HIGHER ORDER H^1 and $H(\nabla\wedge)$ FEM TECHNIQUES with EM APPLICATIONS

DAVOOD ANSARI OGHOL BEIG

A THESIS SUBMITTED FOR THE DEGREE OF
PHD OF ELECTRICAL ENGINEERING

DEPARTMENT OF
ELECTRICAL & COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

February 1, 2011

To the melody of my life, Naghmeh.

Acknowledgments

Honestly, the list of those who have contributed to the progress of this work is tremendously long and I have to rely on my memory to compile this piece of appreciation. Let me put upfront my sincerest apologies to those whom have been dropped this acknowledgment.

First, I shall begin with Professor Leong Mook Seng, my main supervisor at NUS for his support and Patience. I have deep appreciations for my co-adviser Professor Ooi Ban Leong, who was probably the most patient person I ever met and the loss of whom in 2008 caused all of us to suffer.

I should also thank Professor Lee Le-Wei and Alexander Popov from my proposal jury for all their supportive and constructive suggestions. I am deeply grateful to Professors Pavel Solin, A. F. Peterson, Mark Taylor, Tim Warburton, Jan Hesthaven and Leszek Demkowicz with whom I shared my question through email. I also have to express my thankfulness to the members of the *Trilinos* Project at Sandia Labs for their great open-source matrix solver and in this regard I would have to say special thanks to Dr. Christopher Baker and Dr. Michael Heroux for their supportive email guidance. NUS classmates, all of them, and specially Dr. Ng Tiong Huat, Dr. Neelakantam V. Venkatrayalu , Dr. Krishna Agarwal, Dr. Li Ya Nan and Dr. Alexander Shapeev each contributed to my progress in a different sense.

During the last year of my PhD candidature, Professor Jin-Fa Lee offered the opportunity of joining his group as a visiting scholar at OSU's ElectroScience Laboratory where I managed to finish the last part of this work. This has been a life changing experience and I have learned so

much since I joined his group.

Staying away from family and home has not been easy since I left for my PhD in 2005. If there is one thing in this world that made it possible it was the company of my wife, Naghmeh, with her I have shared all the tough moments and indeed the pleasant ones. Last but not least, I shall express thankfulness and gratitude to my parents to meet whom I am running impatient.

Columbus Ohio, February 1, 2011 Davood Ansari Oghol Beig

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction to FEM

Finite element methods (FEMs) form one of the most well funded branches of the numerical techniques developed for solution of integro-differential equations. 'Perhaps, no other family of approximation methods have had a greater impact on the theory and the practice of numerical methods during the twentieth century' [3]. Since the earliest days of their development, FEM's reputation as a powerful numerical method for the solution of integro-differential equations has been constantly growing. Today, FEMs are robustly used in many of the areas once they were expected to have limited capabilities: FEMs were introduced to the engineering society in the 40's but FEMs were not primarily as a good candidate for the solution of problems in electrodynamics, wave phenomenon and fluid dynamics. However, nowadays there are commercial *finite element (FE)* software available in such areas: HFSS, ABAQUS, ANSYS LS-DYNA etc. Expectedly, there is ongoing research involved with development and improvement of FEMs.

In recent decades, the explosive growth of computer capabilities has had a prominent impact on the trend of numerical methods including FEM. The development of *domain decomposition methods (DDMs)* and *parallel processing* scenarios is a natural endeavor for exploitation of the made-available inexpensive parallel computers. In response to the growing demand for compu-

tational accuracy and efficiency, *higher order (HO)* FEMs have received widespread attention [4–16] as a possible remedy for FE dispersion error and geometrical modeling error that become problematic in large scattering problems and over problems with complicated geometries. On the other hand, as FEM techniques become more robust and as technology evolves into areas such as nanophysics, the solution of problems with multi-physical interactions become more and more demanded while the need for improved accuracy/cost rates remains first priority, as ever. With these backgrounds, the research hot fields related to FEMs can be categorized as follows:

1. Development of FEMs for problems with multi-physical complexities.
2. Mesh truncation techniques for *exteriority* problems, i.e. *infinite elements (IEs)* , PML and higher order ABCs.
3. Performance improvements through dedicated matrix solvers and preconditioners, HO / improved basis functions, curvilinear elements etc.
4. Accurate modeling of functional spaces and appropriate treatment of the null-space.
5. Efficient parallelization by means of DDMs and parallel matrix techniques.

1.2 The Scope

The work presented in this dissertation can be divided into the following main streams:

1. Application of semi-optimal nodal sets for improvement of HO *spectral* FE basis functions.
2. Application of optimal nodal sets for improvement of HO spectral IE basis functions.
3. Development of a *dual* -grid based *tree/cotree (T/C)* decomposition for HO null-space treatment of the discrete $H(\nabla\wedge,\Omega)$ space.

4. Development of a *continuously inhomogeneous material property tensor (CIMPT)* approach and associated universal array method for HO FE analysis of problems with continuously varying material properties.

Nevertheless, a great amount of effort has been invested into the realization of a HO FEM software platform that realizes the abovementioned features. Despite the recent developments in the area of HO FEMs, *arbitrary high order (AHO)* spectral FEMs have not been completely extended to H^1 *functional space (H^1)* and (particularly) *curl conforming functional space ($H(\nabla\wedge)$)* problems [17, 18]. While H^1 and $H(\nabla\wedge)$ problems are encountered in many areas of applied physics and enginnering, this work is will mainly deal with examples from the field *electromagnetics (EM)* . On the course of the development of the required software tools, ideas with potnetial novelties will be examined in some of the currently unexplored areas of HO spectral FEMs. Examples include the application of optimal *nodal sets* in FEM analysis of *boundary value problems (BVPs)* and *eigenvalue problems (EVPs)* involving the H^1 and $H(\nabla\wedge)$ functional spaces. HO methods have been applied to the solution of EM problems. Yet, most of the reported works are concerned with *hierarchical* FEs as opposed to spectral elements. Due to their non-spectral nature, such HO implementations cannot benefit from the possible advantages of optimal nodal sets. In other words, there are certain interesting features that can be exploited by opting for spectral but not hierarchical FEs. Practically, application of optimal nodal sets into an AHO H^1 and $H(\nabla\wedge)$ FEM solver will be attempted in this work.

A state-of-the-art FEM platform comprises of wide range of modules, features and technologies. With such a wide variety of features (DDM, multi-physics, HO basis, parallel processing etc.) in mind, a full scale FEM software development calls for the investment of tremendous amounts of manpower. On the other hand, in a research oriented FEM platform development, realization of theoretical and technical innovations relies on the availability of a flexible FEM platform. Such a platform must provide the minimum functionalities defined by the research objectives. Hence, while confining myself to the development of a rather academic FEM platform,

I am mainly concerned with the key aspects that fall in the frameworks of my research interests. Broadly, a spectral HO FEM method will be developed and deployed for problems involving the H^1 and $H(\nabla\wedge)$ functional spaces while certain innovations will be integrated into it. In a last section of this work I also develop the CIMPT approach for HO FE analysis of problems with continuously varying material properties the typical example of which is the ungraded *Luneburg* lens. With the CIMPT method, evaluation of FE matrices becomes much more complicated and hence techniques such as numerical cubatrue must be avoided by all means. In this reagrd, a generalized universal matrix approach is developed which significantly reduces the time required for evaluation of FE matrices.

In brief, I intend to develop an AHO FEM software that effectively handles H^1 and $H(\nabla\wedge)$ problems and integrates the following features:

1. Capability for handling arbitrary element interpolation nodal sets: This allows for the exploitation of the superior interpolatory properties of various available nodal sets. Extensive work has been done on the development of optimal and approximate optimal *precision sets* in *simplex* and *hypercube* shaped elements. However, most related works are engaged with approximation theoretical aspects of optimal nodal sets [19–28]. Other related works have rarely considered simplex shaped elements and seldom attempted extending their approach to $H(\nabla\wedge)$ problems (as opposed to H^1) [29, 30]. Perhaps the closest related works are those of [17, 18] that are limited to H^1 problems and have only considered the *Fekete* nodes while a other nodal sets are nowadays available in the literature. Such interpolation precision sets are expected to deliver performance improvements if integrated into an FEM solver. Once the required machinery for application of arbitrary element interpolation nodal sets (in an FEM solver) is in place, minimal efforts will be needed for examination of any newly reported nodal set. Chapter 4 expands these ideas in detail and adopts an empirical study of the effects of various nodal sets on FE matrix *condition* numbers.

2. Capability for handling exteriority problems: Almost all real world problems are spatially unbounded. A variety of approaches have been proposed for the modeling of infinite domain effects. Among these, IEs [31–41] have made substantial achievements in the recent years. Demkowicz, Rachowicz and Cecot have managed to extend *infinite element methods (IEMs)* into $H(\nabla\wedge)$ vector finite elements concerning EM problems [40, 41]. In comparison to other infinite domain methods such as *boundary element method (BEM)* , IEs are in good harmony with FEMs. The inherent similarity between IEs and FEs simplifies their integration into the structure of a FE engine. Furthermore, due to their similarities with FEMs, some of the technical innovations applicable to FEMs can be extended and used in IEs (including the application of optimal nodal sets). None of the recent developments in IEMs are directly applicable to my case with AHO spectral FEM. Hence as presented in chapter 3, a spectral HO extension of IEs is developed. One of the existing concerns with IE methods is the undesirable growth of matrix condition numbers with the increase in polynomial basis orders[42–44]. Due to their superior interpolatory nature, optimal precision sets are expected to provide means for controlling the mentioned growth in matrix condition numbers. This will be discussed in further detail in chapter 4

3. As an integral part of the FEM code, a generalized universal matrix database had to be developed for evaluation of FE matrices encountered in this work. Luckily, the generalization turned out to be applicable to FEs with curved geometry and non-constant material property tensors. During my visit to Prof. Jin-Fa Lee's group at OSU's ElectroScience Lab., I had the opportunity to integrate this method into one of their 3D DDM codes. This, enabled us to effectively analyze problems where material properties behave as continuously changing functions in space. The results were verified on a non-graded Luneburg lens antenna and form the basis of chapter 5.

The following section presents a classification of a FE software. This classification should by no means be taken as a standard classification of FEMs. Rather, I am intended to provide a typographical identification of my work. At this point, one would see that the realization of the abovementioned features must be integrated into the architectural design of an FEM platform. Some of the critical choices that have been made for the required software development are also addressed in following section. It must be noted that certain features are directly related to the underlying H^1 or $H(\nabla\wedge)$ FE basis. These two types of basis are required for the analysis of two very different types of physical problems.

1.3 Development of Spectral FEM Code

The brief classification of FEMs that is presented in section 1.3.1 is by no means unique, nor it is being provided as an approach for classification of FEMs. The presentation rather serves as a means for classifying the work adopted throughout this research. In section 1.3.2, the major building blocks of a typical FE software platform are introduced; some of the technical details related to the realization of these blocks are also discussed.

1.3.1 Typological Classification

FEMs can be classified in many ways. However, the following key features are often used in the classification of FEMs:

- Time domain versus frequency domain: The distinction between time domain and frequency domain methods is clear as they indicate whether the numerical method is directly implemented in the time domain or otherwise transformed to the frequency domain before actual discretization. Time domain methods are mostly suitable for problems with transient, nonlinear or broadband natures. Very often, frequency domain methods are preferred to time domain methods unless otherwise dictated by the nature of the physical problem,

e.g. nonlinearity a need for wide frequency-spectrum results. This work, however, only concerns with the frequency domain solution of *steady state* time harmonic EM problems.

- Spectral versus hierarchical: FEMs are based on element-wise polynomial expansion of the solution. The expansions can be constructed using either a spectral or hierarchical, basis. In hierarchical FEMs, a basis is used which comprises of polynomials of sucessive orders that satisfy certain orthogonality requirements. The orthogonality helps with the recycling of successively p-refined solutions and construction of certain types of preconditioners. With a hierarchical basis, the resulting solution cannot be directly interpreted as the physical solution and requires some post-processing. Contrarily, when a spectral basis is used, the resulting solution can be directly interpreted as the spectral values of the physical quantity of interest. One of the reasons behind the choice of the spectral basis for this work, is to study the effects of optimal nodal set on the performance of FEs matrices. Nevertheless, the work presented in chapter 5 is based on a hierarchical FE basis.

- h and p adaptivity: h-adaptivity refers to the capability of a FE software for using finer FE mesh to achieve better accuracy whenever required. On the other hand, p-adaptivity refers to the capability of the software in implementing arbitrary order polynomials for the approximation. Together, hp-adaptivity allows FEM to flexibly conform to various geometrical and behavioral complexities. A well engineered FE implementation is expected to be capable of delivering both adaptivity features [45]. The realization of HO spectral FEs reported in this work allows for arbitrary choice of the polynomial order p. Arguably, a possible limitation with such AHO spectral FEM realization is that the polynomial order p accepts a uniform value over the entire mesh. This may often lead to unnecessarily large numbers of global *Degree(s) of Freedom (DoF)* . However, since many of the resulting Degree(s) of Freedom (DoF) are associated with element-internal basis functions, well-developed techniques such as *static condenstaion* [46, 47] can be effectively used to reduce

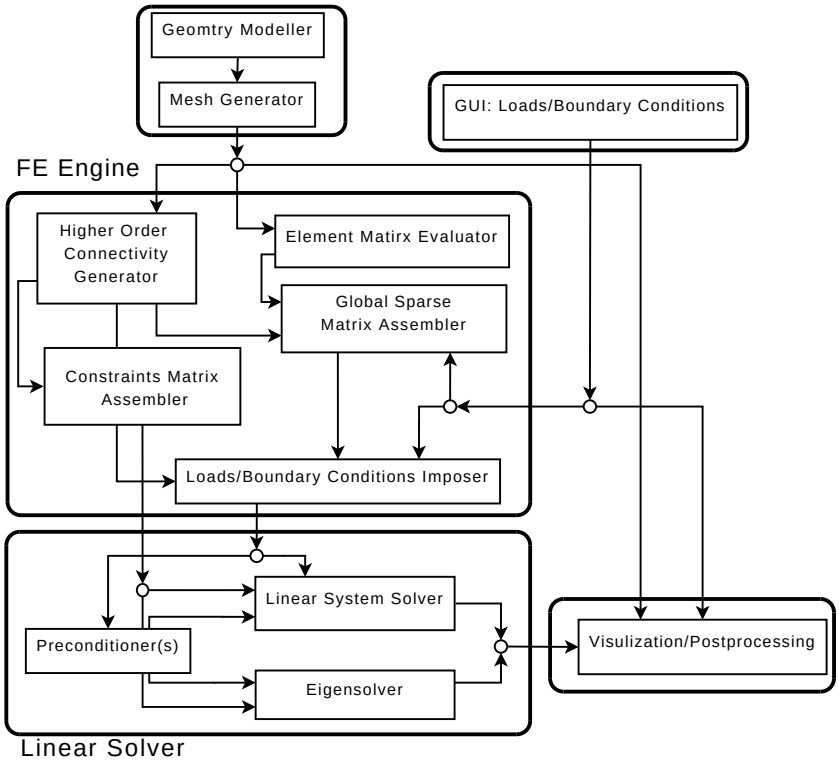


Figure 1.1: The building blocks of a typical FEM platform.

the resulting number of global Degree(s) of Freedom (DoF) without compromising the solution accuracy.

1.3.2 The Necessary Blocks

Depicted in figure 1.1, is the combination of the building blocks of a typical FEM platform. As discussed below, the required blocks are realized in C/C++ with the aid of a variety of third-party library tools.

1. Geometry Model Generator: Most real world problems possess complex geometries. Geometrical modeling is indeed a very active area of applied mathematics with widespread applications in *computer aided design (CAD)* /*computer aided manufacturing (CAM)* engineering. In practice, FEM solution of a problem begins with the construction of a computerized model of the problem geometry. In the absenceWithout of a powerful geometrical modeling tool, FEM's capability in dealing with complex real life applications will

be severely compromised. In this work, geometrical models are constructed using the geometry module of the open-source *Salome* platform. The geometry module is capable of importing/exporting CAD models in a variety of formats such as IGES, STEP and BREP. The module also provides functions for construction of complex geometries.

2. Mesh Generator: Provided with the geometry model, the mesh generator triangulates it into an appropriate set of elements. In this work, the mesh module from the open-source Salome platform is used. Some[1] good open-source mesh generators are available in this module. Moreover, mesh data can be imported/exported in two efficient binary formats: MED and UNV.

3. The FEM Engine: A typical FE engine comprises of the following sub-blocks:

   (a) Element Matrix Evaluator: Technically speaking, this block must provide an accurate and efficient evaluation of individual FE matrices. Some universal matrix methods [48, 49] are extended and realized for this purpose. The required universal matrix libraries are generated using *Mathematica* , *NTL* and *GiNaC* libraries and packed in *HDF5* format. Furthurmore, the widely used *universal* matrix approach is extended to curved elements and elements with CIMPTs. A detailed elaboration of the approach is presented in in chapter 5.

   (b) *Higher order connectivity (HOC)* Generator: Global Degree(s) of Freedom (DoF) connectivity details are necessary for the assemblage of global FE matrices. For HO spectral elements, global element connectivity data can be constructed from first order mesh connectivity data. In my realization, the mesh data from Salome is interfaced into a HOC generator. An efficient *burn and proceed* algorithm (see Appendix A) is designed for the realization of HOC generator.

[1] NETGEN and MEFISTO at the mean time.

(c) Global Sparse Matrix Assembler: A parallel sparse matrix assembler is realized using *Epetra* from the Trilinos [50] package.

(d) Boundary Condition Imposer: Imposition of *Dirichlet* , *Neumann* and *Robin* boundary conditions is a rather straightforward task which is integrated into the FE engine. On the other hand, exteriority boundary conditions, as developed in chapter 3, often require a more thorough treatment. As mentioned earlier in section 1.2, I have chosen to use the IE method for the simulation of infinite domain effects. Refer refer to section 1.2 for the rationale behind the choice of IEM for this research.

(e) Constraints Matrix Assembler: Depending on the problem type, FE matrices can often become *under determined* . In certain cases (as with most EM problems in the $H(\nabla\wedge)$ space), the *rank deficiency* must be identified and treated by imposition of extra constraints [51–60]. For AHO spectral FEM in $H(\nabla\wedge)$, the required constraints can be constructed by explicit imposition of the divergence condition of the *Maxwell* equations [53, 58]. This method, however, requires a *mixed finite element method (MFEM)* formulation that simultaneously involves $H(\nabla\wedge)$ and H^1 spaces and in turn leads to higher computational costs. More efficient approaches have so far been proposed (where MFEM formulations are avoided) for the construction of the constraints[56, 58]. Without some serious extension, such methods cannot be applied to HO spectral FEM as they are either limited to least order elements or rely on the *orthogonal* structure of the hierarchical basis. Hence, as in chapter 2, development of a new method for construction of constraint matrices for HO spectral FEM is presented. The approach utilizes some of the ideas of [58] and exploits the natural properties of the dual FE grid.

4. Linear Solver: The outcome of a FE engine is a linear matrix equation or eigenequation depending on whether the original problem is a BVP or an EVP respectively. Other details

determine if the matrix problem becomes *symmetric* , *hermitian* , real, complex etc. Efficient solution of FE matrix problems relies on the availability of suitable solver mechanisms and appropriate matrix *preconditioners* . In this work, such functionalities are realized in the linear solver block which is developed using the Trilinos package [50]. This is a parallel realization due to the intrinsic parallel nature of Trilinos.

5. Visualization/Post-processing: Visualization is a must for almost any type of scientific data. With HO spectral elements, custom designed visualization tools are needed for proper visualization of FE solutions, FE mesh and Degree(s) of Freedom (DoF). Thus a custom designed visualization block is developed that meets the aforementioned requirements. The block is constructed with the aid of the *Visualization Toolkit* .

1.4 Contributions and Outline

An arbitrary HO FE code was developed which allowed for examination and realization and examination of certain novelties and features related to HO FEMs. Most contributions were achived for both $H(\nabla\wedge)$ and H^1 FEs. The achievements ocan be classified into the followinf four categories:

a. A spectral HO $H(\nabla\wedge)$ conforming FEM was implemented. Most importantly, for the purpose of total elimination of spurious modes, a new T/C decomposition approach was introduced. Unlike its predecessors the method relies on the dual grid and:

   i. It is applicable to arbitrary order spectral basis.

   ii. Global evaluation of discrete path integration and gradient operators is bypassed. The element constraints matrix is geometry independent. Instead, the process is entirely handled using a fixed constraints matrix defined on the reference[2] element.

[2]Often regarded as the master element in the literature.

iii. One-time evaluation of the reference constraints matrix allows for the use of accurate symbolic math methods. Without giving rise to issues with the computational complexity, this results in a more accurate construction of the constraint equations.

iv. It is proved that the constraint equations are solely determined by the topology of the FE mesh.

v. Using the dual *tree* , the sparsity pattern of the constraints matrix is explicitly obtained from the topology of the FE mesh. From efficiency point of view, this allows for preallocation of the sparse storage structure of the matrix and the need for dynamic memory allocation is reduced to a great extent.

Chapter 2 is built on the basis of a detailed discussion of what is itemized in the above. Also, a paper has been published [1] in regards to the above developments.

b. In chapter 3, a HO spectral *finite/infinite element (F/IE)* code was developed for analysis of exteriority problems. A major known issue with HO IEs is the ill-conditioning of the IE matrices. It was shown that the use of semi-optimal interpolation nodal sets can significantly improve the condition numbers of IE matrices. The effect of the choice of Semi-optimal interpolation nodal sets was theoretically and empirically studied in chapter 4. It is shown that the condition number of the *Vandermonde* matrix puts an upper bound on the condition numbers of FE matrices. However, as in the case of $H(\nabla\wedge)$ elements, it was also observed that the condition number of FE matrices is not the only decisive factor in the choice of FE basis or interpolation nodal set. This is because the condition number of the constraints matrix (required for T/C decomposition) is also influenced by the choice of the interpolation nodal set or FE basis. It was concluded that the use of semi-optimal nodal sets is beneficial to the conditioning of FE matrices. It was also concluded that there is a need for the development of nodal sets specific to $H(\nabla\wedge)$ elements. Based on these results, an article has been adjusted and submitted [61].

b. The introduction of the general CIMPT technique for evaluation of FE matrices is one of the other achievement of this work. The method has been discussed in details in chapter 5. Here, we briefly mention that it is based on the fact that material property tensors need not to be element-wise constant functions. Just like element curvature related *Jacobian* terms, material property tensors are in general functions of spatial dimensions. In this regard, an efficient symmetric universal matrix approach was introduced and examined on an example of a Luneburg lens antenna problem. The fact that material property tensors can be treated as non-constant functions elevates the need for re-meshing in cases such as multiphysical problems where one phenomenon determines (and hence changes) the material properties or the *partial differential equation (PDE)* coefficients for another one. The developments associated to chapter 3 are also submitted for publication[62].

Two other articles were also published along with the evolution of this work. One article, presents a generic study on the feasibility of the use of arbitrary HO FE [63] and the other is related to FEM-based extraction of the *Green* 's function of the *Schrodinger* equation encountered in the the study of channel roughness effects in nano-scale semiconductor devices[64].

Chapter 2

Higher Order Spectral FEM for $H(\nabla\wedge,\Omega)$ Problems

---

2.1 Introduction

In this chapter, a new T/C decomposition for $H(\nabla\wedge)$ HO spectral FEs is introduced. The presented approach is a novel extension of the zeroth-order T/C decomposition for HO interpolatory elements which constructs the constraints operator required for elimination of spurious solutions. Earlier works explicitly enforce the divergence condition that requires a mixed FE formulation with both H^1 and $H(\nabla\wedge)$ expansions and involves repeated solutions of *Poisson* 's equation. A recent approach, that avoids the mixed formulation and the Poisson's problem, uses T/C decomposition of edge Degree(s) of Freedom (DoF) over the primal graph and construction of integration and gradient matrices. The approach is easily applied to first order hierarchical elements but becomes quite complex for HO (spectral) elements; In the presence of internal Degree(s) of Freedom (DoF), it is difficult to utilize the primal graph for an explicit decomposition of the spectral Degree(s) of Freedom (DoF). In contrast, this work utilizes the dual grid, resulting in an explicit decomposition of Degree(s) of Freedom (DoF) and construction of constraint equations from a

fixed element matrix. Thus, mixed formulation and Poisson's problems are avoided while eliminating the need for evaluation of integration and gradient matrices. The proposed constraints matrix is element-geometry independent and possesses an explicit sparsity formulation reducing the need for dynamic memory allocation. Numerical examples are included for verification.

2.1.1 Background and Notation

2.1.1.1 Notation

Brought forward, in Tables 2.1 to 2.4, is a list of the symbols and notations used throughout the chapter. The relevance and context of these quantities shall be clear while reading as they appear in the prose. Should the reader encounter any new notation or symbols throughout the article, they may refer to Table 2.1 through Table 2.4 for clarification.

Symbol	Meaning	Dimensions/Units
ϵ_0	free space electric permittivity	farads m^{-1},
μ_0	free space magnetic permeability	Henries m^{-1}
ϵ_r	relative electric permittivity tensor	
μ_r	relative magnetic permeability tensor	
ω	angular frequency (radians per second)	s^{-1}
$\mathcal{E},\mathcal{H}$	electric and magnetic field intensity resp	Voltm^{-1}, Amp.m^{-1}
$\mathcal{A}$	magnetic vector potential	Volt Second m^{-1}
$\mathcal{J}$	electric current density	Amp. m^{d-1}
d	number of physical spatial dimensions	
m	number of disjoint *perfect electric conductors (PECs)*	

Table 2.1: Physical symbols and notations

2.1.1.2 Background

Consider the frequency domain solution of Maxwell's equation (see (2.1)) in a source-free region $\boldsymbol{\Omega}$.

$$\begin{cases} \nabla\wedge\mathcal{E} = -\jmath\omega\mu_0\mu_r\mathcal{H}, & \nabla.\epsilon_0\epsilon_r \quad \mathcal{E} \quad = 0 \\ \nabla\wedge\mathcal{H} = +\jmath\omega\epsilon_0\epsilon_r\mathcal{E}, & \nabla.\mu_0\mu_r \quad \mathcal{H} \quad = 0 \end{cases} \tag{2.1}$$

Symbol	Meaning	Dimensions
$\jmath$	$\sqrt{1}$	
$\triangleq$	define	
$\wedge$	vector wedge product	
cte	constant	
$\ker(\mathcal{L},\boldsymbol{\Xi})$	kernel of linear operator $\mathcal{L}:\boldsymbol{\Xi}\to\boldsymbol{\Theta}$ in $\boldsymbol{\Xi}$	
$\mathcal{L}$	a linear operator between vector spaces, i.e. $\mathcal{L}:\boldsymbol{\Xi}\to\boldsymbol{\Theta}$	
$\cong$	isomorphism of vector spaces	
$\boldsymbol{\Xi}/\boldsymbol{\Theta}$	quotient vector space (when explicitly specified)	
δ_{uv}	*Kronecker* delta function	
$\oplus$	vector space summation sign	
∇_x	the del operator $[\partial_{x_1}\partial_{x_2}\dots\partial_{x_d}]^T$ in the global coordinates	d
∇_ξ	the del operator $[\partial_{\xi_1}\partial_{\xi_2}\dots\partial_{\xi_d}]^T$ in the local coordinates	d
$\boldsymbol{\Omega}$	physical domain of computation, $\boldsymbol{\Omega}\subset\mathbb{R}^d$	
$\mathcal{T},\mathcal{T}(\boldsymbol{\Omega})$	a triangulation of the computational domain $\boldsymbol{\Omega}$	
$\mathcal{G}$	the triangulation $\mathcal{T}$ as a graph	
$\langle\bullet,\bullet\rangle$	*Hilbert* space inner product sign	
G, Q	discrete gradient and path integration operators resp.	
$\vec{F}$	coordinate transform $\vec{F}:\xi\to x$	d
$\vec{\gamma}(t)$	a parametric curve in ξ coordinates	d
x	physical coordinates	d
ξ	reference element coordinates	d
$\triangle$	the reference element	
θ	a point on $\triangle$	
$\{\xi_i\}$	barycentric coordinates on triangle $\triangle$	$d+1$
$\{x_i\}$	global coordinates in $\boldsymbol{\Omega}$	d
J	Jacobian $\mathsf{J}=[\frac{\partial x_i}{\partial\xi_j}]$ for $\vec{F}:\xi\to x$	$d\times d$
K	Jacobian $\mathsf{K}=[\frac{\partial\xi_i}{\partial x_j}]$ for $\vec{F}^{-1}:x\to\xi$	$d\times d$

Table 2.2: Mathematical symbols and notations

In practice, the solution to this system is sought by solving (2.2), which is essentially derived from the curl equations in (2.1).

$$\nabla\wedge\mu_r^{-1}\nabla\wedge\mathcal{E}-\kappa^2\epsilon_r\mathcal{E}=0,\quad \kappa^2\triangleq\epsilon_0\mu_0\omega^2 \tag{2.2}$$

When $\omega=0$, the four equations become independent and the complete solution to (2.1) cannot be sought simply by solving (2.2). By taking the divergence of (2.2), it is observed that it does not enforce the divergence terms of Maxwell equations at $\omega=0$. Hence, when $\omega=0$, solving (2.2) leaves the divergence terms of (2.1) unenforced. In other words, for a given physical domain $\boldsymbol{\Omega}$ and any sufficiently smooth function $\phi:\boldsymbol{\Omega}\to\mathbb{R}$, $\nabla\phi$ turns out to satisfy (2.2) for

Symbol	Meaning	Dimensions
$\mathcal{R}^p(\triangle)$	p-order isometric interpolation nodes in $\triangle$ defined as: $\{\theta\mid\ \xi\|_\theta = \frac{1}{n}(i,j,k), i,j,k \geq 0, i+j+k=n\}$ where ξ is the *barycentric* coordinates representation of points	$\binom{p+d}{d}$
$\mathcal{M}^p(\triangle)$	$\{\theta \in \mathcal{R}^p\mid\ \prod_{i=1}^{d+1} \xi_i\|_\theta \neq 0\}$	$\binom{p-1}{d}$
$\mathcal{B}_i^p(\triangle)$	$\{\theta \in \mathcal{R}^p\mid\ \xi_i\|_\theta = 0\ and\ \forall j \neq i, \xi_j\|_\theta \neq 0\}$	$\binom{p-1}{d-1}$
$\mathcal{N}_i^p(\triangle)$	$\mathcal{M}^p(\triangle)\bigcup\mathcal{B}_i^p(\triangle)$	$\binom{p}{d}$
E	number of elements or faces in $\mathcal{T}$	
N	number of vertices in $\mathcal{G}$	
e	number of edges in $\mathcal{G}$	
e_t	number of tree edges in $\mathcal{G}$	
e_c	number of *cotree* edges in $\mathcal{G}$	

Table 2.3: Nodal precision sets and Topological parameters

$\kappa = 0$ without satisfying the divergence terms of (2.1) (since $\nabla \wedge \nabla\phi = 0$). An immediate observation at this point is that such solutions can be identified as $\ker(\nabla\wedge, \mathbf{V})$, or the kernel of the curl operator in the space of admissible functions $\mathbf{V}$, defined in (2.3).

$$\mathbf{K} = \ker(\nabla\wedge, \mathbf{V}) = \{x \in \mathbf{V} | \nabla \wedge x = 0_{\mathbf{V}}\} \tag{2.3}$$

If the weak form of the EM problem is solely driven by (2.2), such unwanted solutions are directly introduced into FEM results.

2.1.2 Introduction

Despite the introduction of *Nedelec* 's [65, 66] basis to EM FE analysis, a certain class of non-physical solutions (as discussed above) persist in polluting the FEM solutions of EM problems. Nedelec's $H(\nabla\wedge, \mathbf{\Omega})$ conforming elements construct a functional space that can be decomposed into two mutually orthogonal subspaces: Gradient fields and solenoidal fields. The presence of pure gradient fields gives rise to a proper $(\nabla\wedge)$-operator null-space which can lead to nonphysical solutions for *Maxwell* equations. The zero-frequency or static nature of these solutions makes them evanescent to the far field observer. Yet, in the near field EM FE analysis, such zero eigen-valued solutions directly affect the performance of matrix solvers. In microwave engineering literature this problem is regarded as the 'low frequency instability'. A more detailed discussion

Symbol	Meaning	Dimensions
$H^1(\mathbf{\Omega})$	*Sobolev* space $\{f \in L^2(\mathbf{\Omega}) \vert \partial_{x_i} f \in L^2(\mathbf{\Omega})\}$	infinite, cardinal ∮
$H(\nabla\wedge,\mathbf{\Omega})$	Sobolev space $\{f \in (L^2)^d(\mathbf{\Omega}) \vert \nabla \wedge f \in (L^2(\mathbf{\Omega}))^d\}$	infinite, cardinal ∮
$\mathbf{V}$	admissible functions' space.	
	in this case V equals $H(\nabla\wedge,\mathbf{\Omega})$	infinite, cardinal ∮
$\mathbf{K}$	$\ker(\nabla\wedge,\mathbf{V})$ or the functional space of pure gradients	infinite, cardinal ∮
$\mathbf{K}^\perp$	orthogonal complement of $\mathbf{K}$ in $\mathbf{V}$	$\dim \mathbf{V} - \dim \mathbf{K}$
$\mathbf{V}^p$	order p FE discrete analogues of $\mathbf{V}$	$(p+1)(e+e_c p)$
$\mathbf{K}^p$	order p FE discrete analogues of $\mathbf{K}$	$p(e+e_c(p-1)/2!)+e_t$
$\mathbf{K}^{p\perp}$	order p FE discrete analogues of $\mathbf{K}^\perp$	$e_c(p+1)(p+2)/2!$
$\mathbf{\Pi}^p$	order p FE discrete analogues of H^1	$e\,p + E\,p(p-1)/2!+N$
$\mathbf{K}_N$	$\mathbf{K}_N \subset \mathbf{K}$, space of nonphysical gradients (PEC)	$\dim \mathbf{K} - (m-1)$
$\mathbf{K}_N^\perp$	$\mathbf{K}_N^\perp \subset \mathbf{K}$, orthogonal complement of $\mathbf{K}_N$ in $\mathbf{K}$ (PEC)	$m-1$
$\mathbf{K}_P$	$\mathbf{K}_N^\perp \subset \mathbf{K}$, space of physical gradients (PEC)	$m-1$
$\mathbf{K}_N^p$	order p FE discrete analogues of $\mathbf{K}_N$	$\dim \mathbf{K}^p - (m-1)$
$\mathbf{K}_N^{p\perp}$	order p FE discrete analogues of $\mathbf{K}_N^\perp$	$m-1$
$\mathbf{K}_P^p$	order p FE discrete analogues of $\mathbf{K}_P$	$m-1$
$\mathbf{P}^p(\triangle)$	order p $H(\nabla\wedge)$ conforming polynomial space on $\triangle$	$\sum_{i=1}^{d} i\binom{p+1}{i}\binom{d+1}{i+1}$
$P^p(\triangle)$	order p H^1 conforming polynomial space on $\triangle$	$\binom{p+d}{d}$
$\{\alpha_i^{p+1}\}$	canonical basis for $P^{p+1}(\triangle)$, a *Lagrangian* basis defined on $\triangle$ with precision set $\mathcal{R}^{p+1}(\triangle)$	$\dim P^{p+1}(\triangle)$
$\{\boldsymbol{\beta}_i^p\}$	canonical basis for $\mathbf{P}^p(\triangle)$	$\dim \mathbf{P}^p(\triangle)$

Table 2.4: Functional spaces

on such *spurious* modes can be found in [52, 56, 58]. In this work, however, we intend to introduce a novel T/C method that will be used for identification and elimination of the mentioned undesired null-space components.

As discussed in section 2.1.1, direct derivation of the problem weak form from (2.2) results in introduction of spurious solutions. Nevertheless, such spurious solutions can be avoided, either by enforcing the divergence condition(s) [56, 58, 67, 68] or by discarding the $\nabla\phi$ forms from the FEM space of admissible functions. Earlier methods [67, 68] are based on a FE formulation involving in both H^1 and $H(\nabla\wedge)$ FE matrices and requiring repeated solution of an auxiliary Poisson's problem. The authors of [58] revolutionize the approach by avoiding explicit involvement of H^1 matrices and eliminating the Poisson's problem from the matrix solution process. Instead, taking advantage of a primal graph T/C decomposition, they decompose the edge Degree(s) of Freedom (DoF) into the tree and the cotree Degree(s) of Freedom (DoF) and proceed by construction of the so called discrete path integration and gradient matrices. Despite their

global nature, these matrices are entirely composed of $\{0,\pm 1\}$ entries and thus add little computational burden to the solution process. The approach of [58] has been successfully extended to HO hierarchical elements since pure gradient terms of HO hierarchical basis can be discarded at the element level [56].

The approach of [58] can hardly be extended to HO spectral elements since it yields no clear definition for tree and cotree Degree(s) of Freedom (DoF) (specially for internal Degree(s) of Freedom (DoF)). Hence, a direct comparison between this work and that of [58] is not meaningful unless if we confine ourselves to the least order basis. Yet, if we persist on comparing the methods for the least order case, we would observe that they result in identical constraints matrices and identical implementations at the *Krylov* matrix solver level. At this level, [58] deals with a linear operation (which can be formulated as a matrix operation) imposed on the *right hand side (RHS)* of the unconstrained matrix equation, i.e. the *Lanczos* vector. In our approach, the RHS of the matrix equation is premultiplied (once) by a projection operator that confines the RHS to the desired 'gradient free' subspace. Nevertheless, the dual-grid based approach presented in this work has the advantage of bypassing the construction of discrete gradient and integration operators and directly assembling the constraints matrix from a fixed (element geometry invariant) matrix. Note that, with the introduction of HO spectral terms, the generalized version of the so called path integration and gradient matrices of [58] will no longer comprise of simple $\{0,\pm 1\}$ entries. The contributions in the current work can be summarized as follows:

1. It is applicable to arbitrary order spectral basis.

2. It is based on the dual grid (dual to the finite element mesh as a graph).

3. Global evaluation of discrete path integration and gradient operators is bypassed. The element constraints matrix is geometry independent. Instead, the process is entirely handled using a fixed constraints matrix defined on the reference[1] element.

[1] Often regarded as the master element in the literature.

4. One-time evaluation of the reference constraints matrix allows for accurate symbolic math methods to be used. This, without giving rise to issues in computational efficiency, results in a more accurate construction of the constraint equations.

5. It is proved that the constraint equations are solely determined by the topology of the FE mesh.

6. Using the dual tree, the sparsity pattern of the constraints matrix is explicitly calculated. From efficiency point of view, this allows for preallocation of the sparse storage structure of the matrix and eliminates the need for dynamic memory allocation is reduced to a great extent.

Indeed, with minor adjustments, the presented approach can be applied to hierarchical elements since for HO hierarchical elements, the T/C splitting only involves the first order terms of the edge basis [56].

The outline of the article is as follows. Section 2.2 provides the mathematical statement of the null-space problem and the essential idea behind the presented work. Section 2.3 develops a consistent formulation of the involved polynomial spaces and proceeds to present the proposed solution for the single element case. Section 2.4 extends the method into the global case. In presence of multiple disjoint PEC conditions, the physical solution to (2.2) must be allowed to include certain pure gradient functions and thus a complete removal of the pure gradient subspace is undesirable. Hence, as in section 2.5, the constraint equations are reduced so that multiple disjoint PEC cases are covered as well. Section 2.6 derives the explicit formulation for the sparsity of the constraints matrix and section 2.7 shortly discusses some of the aspects encountered during parallel implementation of the proposed dual-grid based method. Section 2.8 briefly discusses the feasibility of extending the method into *three dimensional/dimensions (3D)* . The subsequent sections follow with numerical examples and a final conclusion.

2.2 Vector Space Formulation of the Problem

This section aims at providing the necessary means for identifying the space of spurious solutions in terms of mutually orthogonal vector spaces. As pointed out in section 2.1.1, the problematic solutions of (2.2) share the common property of being associated with zero eigenvalues, i.e. $\kappa = 0$. Assume that the space of admissible functions is denoted by $\mathbf{V}$ and that $\mathbf{K}$ defined by (2.3) is a proper subspace of $\mathbf{V}$. Thus, when $\kappa = 0$, (2.2) will have non-unique solutions in $\mathbf{V}$. Yet, equivalent classes of $\kappa = 0$ solutions are identified as members of the quotient space $\mathbf{V}/\mathbf{K}$ having an isomorphism[2] of the form $\nabla \wedge \mathbf{V} \cong \mathbf{V}/\mathbf{K}$. Further, if $\mathbf{V}$ is a closed inner product space, it is decomposable into a pair of mutually orthogonal (complement) subspaces $\mathbf{K}$ and $\mathbf{K}^{\perp}$ [69, 70] where an isomorphism between $\mathbf{V}/\mathbf{K}$ and $\mathbf{K}^{\perp}$ exist ($\mathbf{V}/\mathbf{K} \cong \mathbf{K}^{\perp}$), see (2.5). Here, the orthogonality, as in (2.4), is defined with respect to the inner product of the Hilbert space $\mathbf{V}$.

$$\mathbf{K}^{\perp} = \{x \in \mathbf{V} | \forall y \in \mathbf{K}, \langle x, y \rangle = 0\} \tag{2.4}$$

$$\mathbf{V} = H(\nabla\wedge, \mathbf{\Omega}) = \mathbf{K} \oplus \mathbf{K}^{\perp} \tag{2.5}$$

At the discrete level, such a decomposition of the space of admissible FEM functions is already given by [71, 72]. It must be emphasized here that the concept not only applies to $\kappa = 0$ solutions of (2.2), but is equally applicable to the solutions of *magnetostatics* problems where the solution of $\nabla \wedge \nabla \wedge A = J$ is sought; see section 2.9.3 for more details.

Our prime intention is to keep the FEM solutions free from unwanted contributions in $\mathbf{K}$. Using [71, 72], it is understood that over every proper finite element triangulation $\mathcal{T}$ of the problem domain $\mathbf{\Omega}$, $\mathbf{K}$ can be constructed from $H^1(\mathbf{\Omega})$ under the effect of the gradient operator: Given a triangulation $\mathcal{T}$ over $\mathbf{\Omega}$, let $\mathbf{V}^p \subset H(\nabla\wedge, \mathbf{\Omega})$ denote the discrete polynomial space of func-

[2] Different notations have been so far used to denote 'isomorphisms'. Readers may encounter notations other that the $\cong$ sign in the literature.

tions constructed by the p-order edge elements [12]. Also, assume that another polynomial space $\mathbf{\Pi}^{p+1}(\mathbf{\Omega}) \subset H^1(\mathbf{\Omega})$, is constructed over the same triangulation $\mathcal{T}$ using $(p+1)$-order scalar elements [14, 73]. Then, the validity of (2.6) is granted where $\mathbf{K}^p$ and $\mathbf{K}^{p\perp}$ in $\mathbf{V}^p$ are the analogous of $\mathbf{K}$ and $\mathbf{K}^{\perp}$ in $\mathbf{V}$. In [52], this has been algebraically illustrated for 0^{th} and 1^{st} order cases.

$$\nabla\mathbf{\Pi}^{p+1} \triangleq \{x|x = \nabla\phi, \phi \in \mathbf{\Pi}^{p+1}\} = \mathbf{K}^p \tag{2.6}$$

The fundamental concept used for the identification of $\mathbf{K}^p$ in $\mathbf{V}^p$ is to find the forward and backward linear transformations between $\mathbf{V}^p$ and $\mathbf{\Pi}^{p+1}$. Due to their nature, the corresponding linear mappings are regarded here as 'gradient' and 'path integration' operators and denoted by $\mathsf{G} : \mathbf{\Pi}^{p+1} \rightarrow \mathbf{V}^p$ and $\mathsf{Q} : \mathbf{V}^p \rightarrow \mathbf{\Pi}^{p+1}$ respectively. Although the transformations have been developed in [58], it must be noted that in [58]:

1. The transformations are directly developed at the global level.
2. The transformations are limited to the least order polynomial basis.
3. The transformations are developed for a different purpose, i.e. explicit imposition of divergence conditions.
4. The subsequent fast implementations are based on the nice properties of first order polynomial basis, i.e. first order polynomial basis the resulting G and Q are made of simple $\{0, \pm 1\}$ entries only but this property does not generalize to HO cases.

2.3 On a Single Element

This section covers the following matters:

- developing the required scalar and vector polynomial bases over a single element.
- introducing the path integration and gradient operators within the element.

- building the constraint equations and the consequent projection operators within a single element.

All of the calculations discussed in the current section were performed using symbolic math software. In other words, all calculations (except for general *curvilinear* elements) discussed in this section can be and should preferably be evaluated by means of analytical methods.

2.3.1 Construction of $H(\nabla\wedge,\Omega)$ Basis

A practical FE mesh comprises of large numbers of elements mostly of *rectilinear* and often of curvilinear type. In the case of rectilinear elements, due to the *affine* nature of the associated coordinate transforms, the mathematical expressions for FE matrices can be *pulled-back* and evaluated (often analytically) over the reference element. With rectilinear elements, the pulled-back expressions are computationally simplified to linear combinations of metric invariant matrices known as universal FEM matrices [48, 49, 74]. Nevertheless, due to the nonlinear nature of curvilinear transforms, the universal matrix approach cannot be extended to arbitrarily curved elements. In this article, however, we shall prove that the proposed constraints matrix is metric independent. In other words, even with curvilinear elements the constraint matrix is entirely determined by the topology of the FE mesh. The proofs, as it will appear in the following sections, involve in a *pull-back* approach.

Few critical points must be clarified here before we engage with the construction of the polynomial basis:

- It is assumed that the physical element is obtained as a transformation (see figure 2.1) of a reference element $\triangle$ in the barycentric coordinates $\{\xi_i\}$. From time to time, $\{\xi_i\}$ are treated as functions of $\{x_i\}$ and vice versa. This should not constitute in any ambiguity since the coordinates transformations are, by definition, *bijective* mappings defined between differentiable *manifolds* .

- The polynomial space is always constructed over the reference element and in terms of the $\{\xi_i\}$ coordinates. Thus, the scalar part of the basis, i.e. the part that specifies the magnitude (and not the direction if any), is always a polynomial function of $\{\xi_i\}$. The only part of the vector basis that is not necessarily a polynomial in ξ coordinates comprises of Ω_{uv} terms of (2.7). This will be discussed in more detail in section 2.3.1.2. The construction of the interpolatory basis is very much similar to that of [12]. The basis is constructed as an extension of the zeroth order-complete basis $\{\Omega_{12}, \Omega_{23}, \Omega_{13}\}$ defined in (2.7).

$$\Omega_{uv} = \xi_u \nabla_x \xi_v - \xi_v \nabla_x \xi_v, \quad 1 \leq u < v \leq d+1 \tag{2.7}$$

  Note that in (2.7), the gradients are defined in the physical coordinates $\{x_i\}$. For the sake of simplicity, we would first assume that the physical element coincides with the reference element. After defining the polynomial spaces for this simplistic case, we shall extend the definitions into the general case.

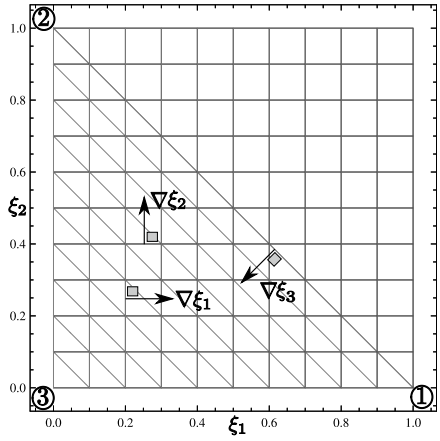
- A formulation similar to (2.7), yields another zero order-complete vector polynomial basis Δ_{uv} directly defined over the reference triangle $\triangle$. The basis components, as defined in (2.8), have geometrical properties over $\triangle$ analogous to those of Ω_{uv} over the physical element.

$$\Delta_{uv} = \xi_u \nabla_\xi \xi_v - \xi_v \nabla_\xi \xi_v, \quad 1 \leq u < v \leq d+1 \tag{2.8}$$

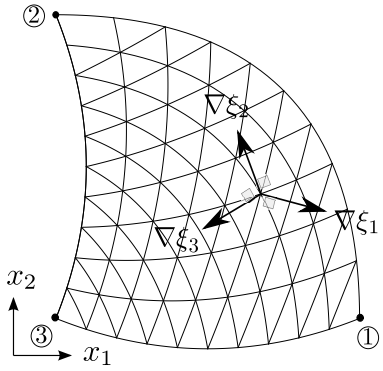
2.3.1.1 Physical Element Coincides with Reference Element

The assumption is that the physical and the reference element are identical and thus the associated coordinates coincide as $\xi_i = x_i$.

As indicated in (2.9) and (2.10), the vector polynomial space $\mathbf{P}^p(\triangle)$ is composed of three



(a) $\triangle = \{(\xi_1, \xi_2) | \xi_1, \xi_2 \geq 0,\ 1 \leq \xi_1 + \xi_2 \leq 1\}$



(b) The physical element (mapped).

Figure 2.1: A visualization of the reference element $\triangle$ in (ξ_1, ξ_2) and the physical element in (x_1, x_2) coordinates. The contours are the plots of $\xi_i = c$ for $c \in \{0, 0.1, \ldots, 1\}$.

linearly dependent subspaces $\mathbf{P}^p_{12}(\triangle)$, $\mathbf{P}^p_{13}(\triangle)$ and $\mathbf{P}^p_{23}(\triangle)$:

$$\mathbf{P}^p(\triangle) = \sum_{1 \leq u < v \leq d+1=3} \mathbf{P}^p_{uv}(\triangle) \tag{2.9}$$

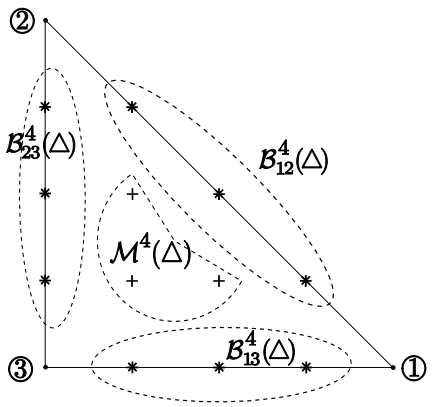
The $\mathbf{P}^p_{uv}(\triangle)$ are in turn defined in terms of three sets of Lagrangian polynomials $\{l^p_{uv,k} | 1 \leq k \leq \frac{(p+1)(p+2)}{2}\}$ defined in (2.11) where θ_i denote points in $\triangle$.

$$\mathbf{P}^p_{uv}(\triangle) = \text{span}(\{\Omega_{uv} l^p_{uv,k} | 1 \leq k \leq \frac{(p+1)(p+2)}{2}\}) \tag{2.10}$$

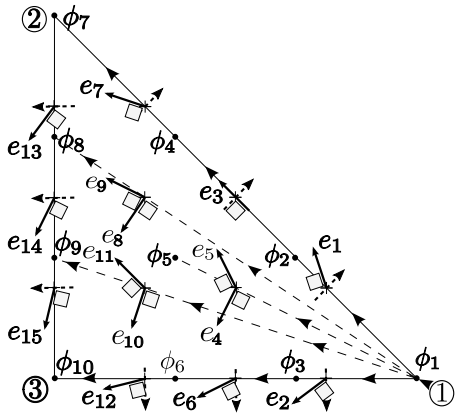
$$l^p_{uv,k}(\xi) \in P^p(\triangle) \text{ such that } \forall \theta_k, \theta_l \in \mathcal{N}^{p+2}_{uv}(\triangle), l^p_{uv,k}(\xi|_{\theta_l}) = \delta_{kl} \tag{2.11}$$

As apparent from (2.11), the three sets of Lagrangians used in the definition of (2.10) span the same polynomial space $P^p(\triangle)$. However, the contrast in their definitions comes form the different precision sets, i.e. $\mathcal{N}^{p+2}_{uv}(\triangle)$, used for their construction in (2.11). For a rigorous definition of $\mathcal{N}^{p+2}_{uv}(\triangle)$ and other related nodal sets refer to Table 2.3. Figure 2.2a depicts $\mathcal{N}^{p+2}_{uv}(\triangle)$, $\mathcal{M}^{p+2}(\triangle)$ and $\mathcal{B}^{p+2}_{uv}(\triangle)$ for $p = 2$.

The interpolatory basis used for the construction of $\mathbf{P}^p(\triangle)$ consists of a pair of polynomials,



(a) Building blocks of the nodal sets used for 2^{nd}-order vector polynomials ($p = 2$). Note that $\mathcal{N}^{p+2}_{uv}(\triangle) = \mathcal{B}^{p+2}_{uv}(\triangle) \bigcup \mathcal{M}^{p+2}(\triangle)$ for $1 \leq u < v \leq d = 2$.



(b) Integration path, vector DoF (direction only) and scalar DoF on the same stencil. $\{\phi_i\}$ and $\{e_i\}$ are the actual DoF in $\mathbf{\Pi}^{p+1}$ and $\mathbf{V}^p$ respectively.

Figure 2.2: 2^{nd}-order nodal sets and DoF (Degrees of Freedom) of $H(\nabla\wedge,\mathbf{\Omega})$ basis on the triangle.

$\Omega_{23}l^p_{23,m}(\xi)$ and $\Omega_{13}l^p_{13,m}(\xi)$, for every internal node $m \in \mathcal{M}^{p+2}(\triangle)$, plus $(p+1)$ polynomials of the form $\Omega_{uv}l^p_{uv,e}(\xi), e \in \mathcal{B}^{p+2}_{uv}(\triangle)$ for each edge 'uv' where $\mathcal{B}^{p+2}_{uv}(\triangle)$ denotes the set of nodes on edge 'uv'. For the $p = 2$ case, the the associated nodal sets have been visualized in figure 2.2a. Exact definitions of these sets can be found in Table 2.3. Now, having the basis sorted in a linear order, $\mathbf{P}^p(\triangle)$ can be expressed as (2.12).

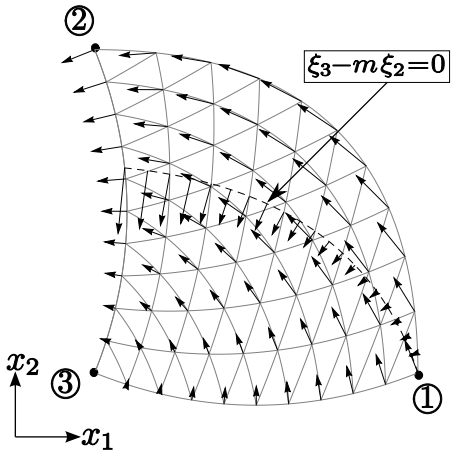
$$\mathbf{P}^p(\triangle) = \text{span}(\{\vec{\beta}^p_k | 1 \leq k \leq \dim \mathbf{P}^p(\triangle)\}), \dim \mathbf{P}^p(\triangle) = (p+1)(p+3) \tag{2.12}$$

For technical reasons, we are also interested to separate the Ω_{uv} part of the basis terms from the pure Lagrangian terms and sort it in the linear order. This is reflected in (2.13) where $\Omega_{u_kv_k}$ and l^p_k respectively point to the Ω_{uv} and Lagrangian terms associated to each basis component.

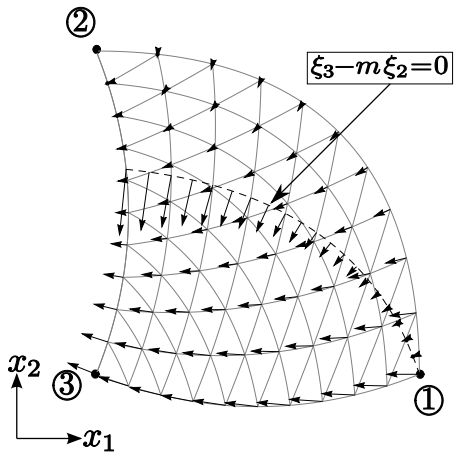
$$\mathbf{P}^p(\triangle) = \text{span}(\{l^p_k \Omega_{u_kv_k} | 1 \leq k \leq \dim \mathbf{P}^p(\triangle)\}), \dim \mathbf{P}^p(\triangle) = (p+1)(p+3) \tag{2.13}$$

Figure 2.2b is a plot of the Degree(s) of Freedom (DoF) associated to the basis (2.12). It can

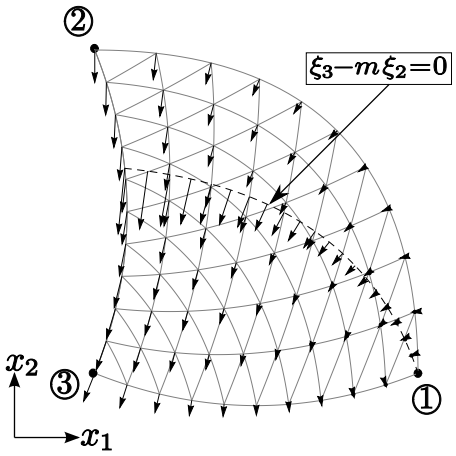
be verified that Ω_{uv}s, exhibit a rotational symmetry causing them to be normal to two of the tree triangle edges (for example see figure 2.3d). Thus, it can be shown that each $\Omega_{uv} l^p_{uv,k}$ has both tangential and normal components on edge 'uv' and solely normal ones on the other two. The mentioned rotational symmetry plays critical role in the T/C decomposition of element Degree(s) of Freedom (DoF) as it will be brought forward in section 2.3.3.



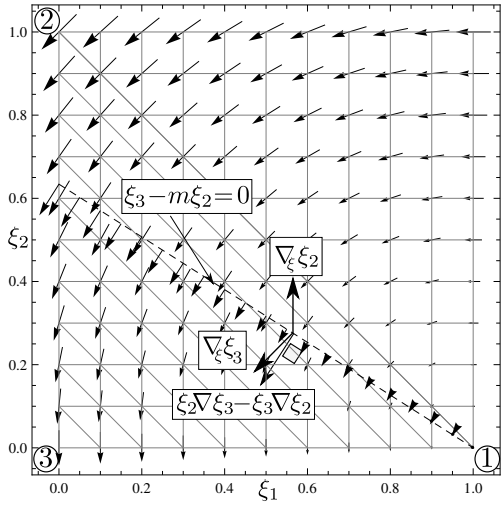
(a) Physical element with the filed plot of Ω_{12} and Ω_{23} over $\xi_3 = m\xi_2$



(b) Physical element with the filed plot of Ω_{13} and Ω_{23} over $\xi_3 = m\xi_2$



(c) Physical element with the filed plot of Ω_{23} and Ω_{23} over $\xi_3 = m\xi_2$



(d) Reference element with field plots (direction only) of $\Delta_{23} = \xi_2\nabla_\xi\xi_3 - \xi_3\nabla_\xi\xi_2$

Figure 2.3: Plots of Ω_{uv} and Δ_{23} respectively over the physical and the reference element emphasizing on the orthogonality of Ω_{23} to $\xi_3 - m\xi_2 = 0$ in the physical element domain and the orthogonality of Δ to $\xi_3 - m\xi_2 = 0$ in $\triangle$.

2.3.1.2 General Case With (Possibly) Curved Elements

The $H(\nabla\wedge)$ basis introduced in the preceding section can be easily extended to the general case with (possibly) curvilinear elements. This can be achieved by having the ξ_is considered as a functions of the physical coordinates x_i. Thus, with a general curvilinear transformation between the reference and the actual physical element, Ω_{uv} of (2.7) would not be confined to polynomial functions in $\{x_i\}$s. Figure 2.3, plots the zero-order edge basis of (2.7) over a curved element. Note at the way Ω_{uv} perpendicularly crosses the edges '$u'v'$' where $\{u', v'\} \neq \{u, v\}$. It is straight forward to see that Ω_{uv} of (2.7) is always perpendicular to edges '$u'v'$' where $\{u', v'\} \neq \{u, v\}$: When $\xi_u = 0$, Ω_{uv} equals $-\xi_v \nabla_x \xi_u$ which is perpendicular to the $\xi_u = 0$ contour in $\{x_i\}$ coordinates. Similarly, when $\xi_v = 0$, Ω_{uv} equals $\xi_u \nabla_x \xi_v$ which is perpendicular to the $\xi_v = 0$ contour.

2.3.2 A Summary of Transformation Rules

In the sections that will follow, we will develop the so called HO discrete gradient and path integration operators. These matrix operators will be used for the construction of the so called constraints matrix. Taking advantage of a pull-back approach we shall prove that the discrete gradient and path integration operators are independent from element shape and metric properties. The pull-back approach involves transformation rules between the physical and the reference element. Here, we would provide a summary of the required transformation rules that would, later on, be used in the mentioned proofs and derivations.

As depicted in figure 2.1, the physical element is obtained by applying a coordinate transformation on the reference element element $\triangle$. Let $\vec{F} : \xi \rightarrow x$ denote the coordinate transformation from the reference element $\triangle$ to the physical element. It is a general rule that in a coordinate system $\{x_i\}$, isosceles $f(x) = cte$ contours are orthogonal to the $\nabla_x f(x)$. For the two coordinate systems of interest, i.e. ξ and x, this can be formulated as:

$$\nabla_x f(x) \perp f(x) = cte \tag{2.14}$$

$$\nabla_\xi f \circ \vec{F}(\xi) \perp f \circ \vec{F}(\xi) = cte \tag{2.15}$$

Let us define the Jacobian matrix associated to $\vec{F}$ as $\mathsf{J} \triangleq [\frac{\partial x_i}{\partial \xi_j}] = [\frac{\partial F(\xi)_i}{\partial \xi_j}]$. For convenience we would also define the inverted Jacobian matrix as $\mathsf{K} = \mathsf{J}^{-1} = [\frac{\partial \xi_i}{\partial x_j}] = [\frac{\partial F^{-1}(x)_i}{\partial x_j}]$. It can be verified that gradients in the two coordinate systems are correlated by the following transformation rules:

$$\nabla_x f(x) = \mathsf{K}^T \nabla_\xi f \circ \vec{F}(\xi) \tag{2.16}$$

$$\nabla_\xi f \circ \vec{F}(\xi) = \mathsf{J}^T \nabla_x f(x) \tag{2.17}$$

We would also like to see how infinitesimal path integration vectors (and similarly tangent vectors) along parametric curves $\xi = \vec{\gamma}(t)$ transform between the two coordinate systems. Infinitesimal path integration vectors $\mathrm{d}\vec{l}_x$ and $\mathrm{d}\vec{l}_\xi$ are here defined in (2.18) and (2.19).

$$\mathrm{d}\vec{l}_x|_{x=\vec{F}\circ\vec{\gamma}(t)} \triangleq [dx_i]\,|_{x=\vec{F}\circ\vec{\gamma}(t)} = \left[\frac{\partial F_i \circ \vec{\gamma}(t)}{\partial t}\right] d\,t \tag{2.18}$$

$$\mathrm{d}\vec{l}_\xi|_{\xi=\vec{\gamma}(t)} \triangleq [d\xi_i]\,|_{\xi=\vec{\gamma}(t)} = \left[\frac{\partial \gamma_i(t)}{\partial t}\right] d\,t \tag{2.19}$$

Simple application of the *chain* rule on (2.18) and (2.19) implies:

$$\mathrm{d}\vec{l}_x|_{x=\vec{F}\circ\vec{\gamma}(t)} = \mathsf{J}\,\mathrm{d}\vec{l}_\xi|_{\xi=\vec{\gamma}(t)} \tag{2.20}$$

$$\mathrm{d}\vec{l}_\xi|_{\xi=\vec{\gamma}(t)} = \mathsf{K}\,\mathrm{d}\vec{l}_x|_{x=\vec{F}\circ\vec{\gamma}(t)} \tag{2.21}$$

Finally, by integration of (2.16), (2.8) and (2.7), it is concluded that Ω_{uv} and Δ_{uv} are related through the following transformation rules:

$$\Omega_{uv} = \xi_u \nabla_x \xi_v - \xi_v \nabla_x \xi_u = \mathsf{K}^T \left(\xi_u \nabla_\xi \xi_v - \xi_v \nabla_\xi \xi_u\right) = \mathsf{K}^T \, \Delta_{uv} \tag{2.22}$$

$$\Delta_{uv} = \mathsf{J}^T \, \Omega_{uv} \tag{2.23}$$

2.3.3 The Path Integration Operator Q

Suppose that a scalar function ϕ is expanded using a Lagrangian H^1 basis in $P^{p+1}(\triangle)$. Also assume that the vector function $\nabla_x \phi$, is expanded using an interpolatory $H(\nabla\wedge,\mathbf{\Omega})$ basis in $\mathbf{P}^p(\triangle)$. A symbolic description of the involved Degree(s) of Freedom (DoF) for the 2^{nd}-order case can be found in figure 2.2b. In (2.24), $\{\alpha_u^{p+1}\}$ are the scalar Lagrangians defined over $\mathcal{R}^{p+1}(\triangle)$.

$$P^{p+1}(\triangle) = \text{span}\{\alpha_u^{p+1} | \alpha_u^{p+1}(n_v) = \delta_{uv}, \forall n_u, n_v \in \mathcal{R}^{p+1}(\triangle)\} \tag{2.24}$$

If ϕ_1 (the value of ϕ at node ①) is taken as the reference, then the remaining scalar Degree(s) of Freedom (DoF), i.e. $\{\phi_2, ..., \phi_{\dim P^{p+1}(\triangle)}\}$, can be expressed in terms of a path integration of the $\nabla_x \phi$ from node ① to the corresponding node. This is mathematically stated in (2.25) where θ_i represent the associated interpolation nodes over the reference element.

$$\phi_u = (\phi_1 = 0) + \int_{\theta_1}^{\theta_u} \nabla_x \phi \cdot \mathrm{d}\vec{l}_x|_{x=\vec{F}\circ\vec{\gamma}(t)}, \quad 1 \leq u \leq \dim P^{p+1} \tag{2.25}$$

$$\vec{\gamma}(t) = \begin{cases} \xi_1 = t \\ \\ \xi_2 = mt \end{cases} \tag{2.26}$$

Here, as formulated in (2.26), the integration path is chosen to be the straight line passing

though node ① and the destination point. However, the integration path would not necessarily be an straight line in the physical domain. For the 2^{nd}-order case, examples of the integration paths which start from node ① are depicted (over the reference element) using dashed lines and thick arrow heads in figure 2.2b. Figure 2.3 visualizes an example of the integration path on both the reference and the physical elements. If $\nabla_x\phi$ in (2.25) is expanded over the basis of (2.13), (2.25) can be restated as (2.27).

$$\phi_u = \int_{\theta_1}^{\theta_u} \sum_k e_k l_k \Omega_k \cdot \mathrm{d}\vec{l}_x|_{x=\vec{F}\circ\vec{\gamma}(t)} \tag{2.27}$$

In (2.27), $\mathrm{d}\vec{l}_x$ is the infinitesimal path integration vector along $\vec{F}\circ\vec{\gamma}(t)$ which according to (2.26) expresses the image of the line passing though node ① of $\triangle$ under the effect of F. Using the transformation rules (2.20) and (2.22), (2.27) can be pulled-back into the reference element and expressed in $\{\xi_i\}$ coordinates. Doing so yields (2.28) which is entirely expressed in terms of reference element related entities.

$$\phi_u = \int_{\theta_1}^{\theta_u} \sum_k e_k l_k (\mathsf{K}^T \Delta_{u_k v_k})^T \mathsf{J}\, \mathrm{d}\vec{l}_\xi|_{\xi=\vec{\gamma}(t)} = \sum_k e_k \int_{\theta_1}^{\theta_u} l_k\, \Delta_{u_k v_k}^T \mathrm{d}\vec{l}_\xi|_{\xi=\vec{\gamma}(t)} \tag{2.28}$$

Equation (2.28) reflects a linear relation Q between the scalar Degree(s) of Freedom (DoF) (2.29) and the vector Degree(s) of Freedom (DoF) (2.30). In this manner, Q shall consist of $\dim P^{p+1}\text{–}1$ nonzero rows, since the row that corresponds to ϕ_1 is all zeros. Two key observations should be made at this point:

$$\{\phi_1 = 0, \phi_2, \ldots, \phi_{\dim P^{p+1}(\triangle)}\} \tag{2.29}$$

$$\{e_1, e_2, \ldots, e_{\dim \mathbf{P}^p(\triangle)}\} \tag{2.30}$$

1. Among the integral terms in (2.28), contributions from $\mathbf{P}^p_{23}(\triangle)$ would vanish. In other words, entries (columns) of Q for in which Δ_{23} is involved would be entirely zero. This is due to the geometric alignment of $\mathbf{P}^p_{23}(\triangle)$ along Ω_{23} which is always perpendicular (see figure 2.2b and figure 2.3c) to the integration paths. Generally speaking, it can be shown that the parametric curve $F\circ\gamma(t)$ of (2.26) is perpendicular to Ω_{23} and its multiplication by any scalar function defined over the physical element (see figure 2.3). This important property of the basis shall be regarded as the rotational symmetry of the basis around node ①. Note that the symmetry is intentional and achieved by design of the basis.

2. As apparent from (2.28), the linear relation Q, which will be regarded as the discrete path integration operator is independent from element metric properties.

Based on the these observations, the vector Degree(s) of Freedom (DoF) in $\mathbf{P}^p(\triangle)$ are partitioned into two groups:

1. Those related to $\mathbf{P}^p_{23}(\triangle)$ or equivalently those for which the associated columns in Q are zero.

2. The complement to those specified in item 1.

In analogy to [52] and [58], the latter shall be regarded as the (local) 'tree' Degree(s) of Freedom (DoF) while the former is called the (local) 'cotree' Degree(s) of Freedom (DoF). Here, the 'local' modifier is used to emphasize on the fact that the mentioned local tree and local cotree Degree(s) of Freedom (DoF) are defined at the element level and should not be confused with global tree and global cotree Degree(s) of Freedom (DoF) introduced later in section 2.4.2. Denoting the partitioning by subscript 't' and 'c' and reordering the vector Degree(s) of Freedom (DoF) of $\mathbf{P}^p(\triangle)$ accordingly, one writes:

$$[\phi_1\dots\phi_{\dim P^{p+1}(\triangle)}]^T=\mathsf{Q}[e_1\dots e_{\dim\mathbf{P}^p(\triangle)}]^T,\ \mathsf{Q}=[\mathsf{Q}_t\mid\mathsf{Q}_c]=[\mathsf{Q}_t\mid 0_c]^T \tag{2.31}$$

Note that both edge and internal Degree(s) of Freedom (DoF) are present among the mentioned tree and cotree Degree(s) of Freedom (DoF). For example, the tree and the cotree Degree(s) of Freedom (DoF) for the case of figure 2.2b can be denoted by (2.32) and (2.33) respectively:

$$\{e_1, e_2, e_3, e_5, e_7, e_9, e_{11}, e_{12}\} \tag{2.32}$$

$$\{e_4, e_8, e_{10}, e_{13}, e_{14}, e_{15}\} \tag{2.33}$$

In general, every p-order $H(\nabla\wedge,\mathbf{\Omega})$ element possesses $(p+1)(p+3)$ Degree(s) of Freedom (DoF) out of which $(p+1)(p+2)/2!$ Degree(s) of Freedom (DoF) (one edge and one half of the internal Degree(s) of Freedom (DoF)) belong to the cotree. On the other hand, there remain $(p+1)(p+4)/2$ Degree(s) of Freedom (DoF) consisting of $2(p+1)$ Degree(s) of Freedom (DoF) for the two non-cotree edges and one Degree(s) of Freedom (DoF) per every internal node ($p(p+1)/2$). Take note that for the zeroth order-complete elements, the tree and cotree Degree(s) of Freedom (DoF) defined here coincide with those of [52, 56, 58].

2.3.4 The Gradient Operator G

The argument presented here is the converse of that of section 2.3.3. Similar to section 2.3.3, assume that the scalar and the vector polynomial spaces are expressed using the associated canonical basis as stated in (2.24) and (2.13) respectively. If ϕ and $\nabla\phi$ are expanded in $P^{p+1}(\triangle)$ and $\mathbf{P}^{p}(\triangle)$ respectively, by matching the two expansions (2.34) is immediately deduced.

$$\sum_{k=1}^{M} e_k\, l_k^p\, \Omega_{u_k v_k} = \sum_{v=1}^{N} \phi_v \nabla_x \alpha_v^{p+1},\ M = \dim \mathbf{P}^p(\triangle),\ N = \dim P^{p+1}(\triangle) \tag{2.34}$$

Equation (2.34) can be tested at various interpolation points $\vec{F}\circ\xi|_{\theta_k}$ (in the physical element) and along various directions $\vec{\tau}_k$ (again in the physical domain). In regards to the choice of the

testing vector $\vec{\tau}_k$ the following two types of situation arise. We shall see that the invariance of the discrete gradient operator under the change of element geometry depends on the choice of the testing vectors.

1. For Degree(s) of Freedom (DoF) e_k, the interpolation node $\vec{F}\circ\xi|_{\theta_k}$ is located on edge 'u_kv_k'. In this situation $\vec{\tau}_k$ is chosen to be $\frac{\partial\vec{F}\circ\vec{\gamma}(t)}{\partial t}|_{\theta_k}$ where we define $\gamma(t)$ as $\{\xi_{u_k} = 1-t, \xi_{v_k} = t\}$. This simply yields a vector tangent (at $\vec{F}\circ\xi|_{\theta_k}$) to the curve that falls on the 'uv' edge of the element. In this manner, all the contributions from $\mathbf{P}^p_{u'v'}(\triangle)$ where $\{uv\} \neq u'v'$ are eliminated as they are perpendicular to the 'uv' edge. Furthermore, except for one term all contributions from $\mathbf{P}^p_{uv}(\triangle)$ are also eliminated due to the interpolatory properties of the Lagrangians involved in $\mathbf{P}^p_{uv}(\triangle)$. Thus, from (2.34) equations of the form (2.35) would result.

2. For pairs of internal Degree(s) of Freedom (DoF) e_k and e_{k+1}, the interpolation node $\vec{F}\circ\xi|_{\theta_k}$ is located inside the element. Note that for internal Degree(s) of Freedom (DoF) two consecutive Degree(s) of Freedom (DoF) e_k and e_{k+1} are defined at each interpolation point. Hence, θ_k and θ_{k+1} are identical points. In this situation, (2.34) must be tested twice at each node. Two testing vectors, say $\vec{\tau}_{k1}$ and $\vec{\tau}_{k2}$, would hence be needed. In this case, the testing vectors are chosen to be $\frac{\partial\vec{F}\circ\vec{\gamma}(t)}{\partial t}|_{\theta_k}$ where $\gamma(t)$ is respectively defined as $\{\xi_1 = t, \xi_2 = \xi_2|_{\theta_k}, \xi_3 = 1-\xi_1-\xi_2\}$ and $\{\xi_2 = t, \xi_1 = \xi_1|_{\theta_k}, \xi_3 = 1-\xi_1-\xi_2\}$. Thus $\vec{\tau}_{k1}$ and $\vec{\tau}_{k2}$ would respectively become the vectors tangent to $\xi_2 = cte$ and $\xi_2 = cte$ contours at $\vec{F}\circ\xi|_{\theta_k}$. Testing (2.34) with $\vec{\tau}_{k1}$ and $\vec{\tau}_{k2}$ at $\vec{F}\circ\xi|_{\theta_k}$ yields two linear equations. Due to the interpolatory properties of the Lagrangians, except for the e_k and e_{k+1} terms, all terms on the *left hand side (LHS)* of (2.34) would vanish and the resulting equations can be lumped into (2.36).

$$e_k\ (l_k^p\ \Omega_{u_kv_k}.\vec{\tau}_k)|_{\vec{F}\circ\xi|_{\theta_k}} = \sum_{v=1}^{N}\phi_v(\nabla_x\alpha_v^{p+1}.\vec{\tau}_k)|_{\vec{F}\circ\xi|_{\theta_k}},\ N=\dim P^{p+1}(\triangle) \tag{2.35}$$

$$\left[\begin{matrix} l_k^p\ \Omega_{u_kv_k}.\vec{\tau}_{k1} & l_{k+1}^p\ \Omega_{u_{k+1}v_{k+1}}.\vec{\tau}_{k1} \\ l_k^p\ \Omega_{u_kv_k}.\vec{\tau}_{k2} & l_{k+1}^p\ \Omega_{u_{k+1}v_{k+1}}.\vec{\tau}_{k2}\end{matrix}\right]\Bigg|_{\vec{F}\circ\xi|_{\theta_k}}\left[\begin{matrix} e_k \\ e_{k+1}\end{matrix}\right]=\sum_{l=1}^{N}\phi_l\left[\begin{matrix}\nabla_x\alpha_l^{p+1}.\vec{\tau}_{k1} \\ \nabla_x\alpha_l^{p+1}.\vec{\tau}_{k2}\end{matrix}\right]\Bigg|_{\vec{F}\circ\xi|_{\theta_k}},N=\dim P^{p+1}(\triangle) \tag{2.36}$$

Let $\vec{t}_k$, $\vec{t}_{k1}$ and $\vec{t}_{k2}$ respectively denote the $\{\xi\}$ coordinate analogues of the $\vec{\tau}_k$, $\vec{\tau}_{k1}$ and $\vec{\tau}_{k2}$. For example, $\vec{t}_k$ denotes $\frac{\partial\vec{\gamma}(t)}{\partial t}|_{\theta_k}$ where $\gamma(t)=\{\xi_{u_k}=1-t,\xi_{v_k}=t\}$. This would be the vector tangent (at $\xi|_{\theta_k}$) to the curve (actually the straight line) that falls on the 'uv' edge of the reference element. It can be verified that these vectors follow the same transformation rules provided in (2.20) and (2.21), e.g. $\vec{t}_k = \mathsf{J}\ \vec{\tau}_k$. Hence, using transformation rules (2.16), (2.20) and (2.22), (2.35) and (2.36) can be pulled-back into the reference element and rephrased as (2.37) and (2.38) respectively.

$$e_k\ (l_k^p\ \Delta_{u_kv_k}.\vec{t}_k)|_{\vec{F}\circ\xi|_{\theta_k}} = \sum_{v=1}^{dimP^{p+1}(\triangle)}\phi_v(\nabla_\xi\alpha_v^{p+1}.\vec{t}_k)|_{\vec{F}\circ\xi|_{\theta_k}} \tag{2.37}$$

$$\left[\begin{matrix} l_k^p\ \Delta_{u_kv_k}.\vec{t}_{k1} & l_{k+1}^p\ \Delta_{u_{k+1}v_{k+1}}.\vec{t}_{k1} \\ l_k^p\ \Delta_{u_kv_k}.\vec{t}_{k2} & l_{k+1}^p\ \Delta_{u_{k+1}v_{k+1}}.\vec{t}_{k2}\end{matrix}\right]\Bigg|_{\xi|_{\theta_k}}\left[\begin{matrix} e_k \\ e_{k+1}\end{matrix}\right]=\sum_{l=1}^{N}\phi_l\left[\begin{matrix}\nabla_\xi\alpha_l^{p+1}.\vec{t}_{k1} \\ \nabla_\xi\alpha_l^{p+1}.\vec{t}_{k2}\end{matrix}\right]\Bigg|_{\xi|_{\theta_k}},N=\dim P^{p+1}(\triangle) \tag{2.38}$$

Equation (2.35) and (2.36) can be solved to yield linear expressions of vector Degree(s) of Freedom (DoF) $\{e_i\}$ in terms of scalar Degree(s) of Freedom (DoF) $\{\phi_i\}$. The resulting linear relation (as is (2.39)) is regarded as the gradient operator G.

$$[e_1 \dots e_{\dim \mathbf{P}^p(\triangle)}]^T = \mathsf{G}\,[\phi_1 \dots \phi_{\dim P^{p+1}(\triangle)}]^T \tag{2.39}$$

It is critical to observe that (2.35) and (2.36) are both expressed in terms entities entirely related to the reference element. Thus, these equations and the resulting G operator are completely independent from the physical element's metric properties. It must be notified here that with other choices of testing vectors $\vec{\tau}_i$ the resulting G operator would not necessarily become insensitive to the metric properties of the physical element.

2.3.5 The Constraints

As depicted in figure 2.4, Q and G are designed to transform row vector representations of polynomials in $\mathbf{P}^p(\triangle)$ and $P^{p+1}(\triangle)$ with respect to the path-integration and gradient operations. Consider that Q and G are the discrete analogous of path integration and gradient operators. If E and Φ are the row vector representation of two polynomials in $\mathbf{P}^p(\triangle)$ and $P^{p+1}(\triangle)$ and they happen to satisfy (2.40) and (2.41) simultaneously, it is concluded that E is the row vector representation of a polynomial in $\nabla P^{p+1}(\triangle) \subset \mathbf{P}^p(\triangle)$.

$$\mathsf{E} = \mathsf{G}\,\Phi \tag{2.40}$$

$$\Phi = \mathsf{Q}\,\mathsf{E} \tag{2.41}$$

Recall from section 2.2, that $\mathbf{K}^p = \nabla\mathbf{\Pi}^{p+1}$ and thus in this case $\mathbf{K}^p = \nabla P^{p+1}(\triangle)$. Hence, following (2.31) and the mentioned T/C partitioning, $\mathbf{K}^p(\triangle)$ can be expressed as a subspace of $\mathbf{V}^p(\triangle) = \mathbf{P}^p(\triangle)$ satisfying (2.42). Take note that E is the row vector representation of polynomials in $\mathbf{P}^p(\triangle)$ with respect to the $\{\vec{\beta}_u^p\}$ basis of (2.12).

$$\begin{aligned}
&\mathsf{E}^T[\beta_1^p \cdots \beta_{(p+1)(p+3)}^p]^T \in \mathbf{K}^p(\triangle) \Longleftrightarrow \left[\begin{array}{c}\mathsf{E}_t\\\hline \mathsf{E}_c\end{array}\right] = \mathcal{C}\left[\begin{array}{c}\mathsf{E}_t\\\hline \mathsf{E}_c\end{array}\right],\\
&\mathcal{C} \triangleq \left[\begin{array}{c|c}\mathsf{G}_t\mathsf{Q}_t & \mathsf{G}_t\mathsf{Q}_c\\\hline \mathsf{G}_c\mathsf{Q}_t & \mathsf{G}_c\mathsf{Q}_c\end{array}\right] = \left[\begin{array}{c|c}\mathsf{I}_{tt} & \mathsf{0}_{tc}\\\hline \mathsf{G}_c\mathsf{Q}_t & \mathsf{0}_{cc}\end{array}\right]
\end{aligned} \tag{2.42}$$

$$\mathsf{E}_c = \mathsf{C}\,\mathsf{E}_t,\ \mathsf{C} \triangleq \mathsf{G}_c\mathsf{Q}_t \tag{2.43}$$

In (2.43), the C matrix is defined as $\mathsf{G}_c\mathsf{Q}_t$. C is regarded as the constraints matrix. Following G and Q, C is independent from element metric properties. In other words, C does not depend on individual element shape and dimensions. Equation (2.44) provides examples of the constraints matrix C for the $p = 1, 2$ cases. Take note that, $\mathcal{C}$ can be constructed solely by knowing C due to the trivial structure of the other blocks in $\mathcal{C}$.

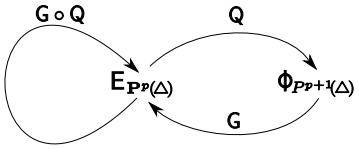


Figure 2.4: Composition of the gradient and the path integration maps.

$$\mathsf{C}_{p=2} = \frac{1}{3}\begin{bmatrix}
-7 & -7 & 2 & 3 & 2 & 0 & 1 & -2 & 0\\
-3 & -3 & -2 & -2 & 2 & 0 & 2 & -2 & 0\\
-3 & -3 & 2 & 2 & -2 & 0 & 1 & -1 & 0\\
-2 & -2 & 0 & -1 & 2 & -3 & 0 & -2 & 0\\
-2 & -2 & 1 & 0 & 1 & 0 & 2 & -4 & 0\\
-2 & -2 & 2 & 1 & 0 & 0 & 1 & 0 & -3
\end{bmatrix}, \mathsf{C}_{p=1} = \frac{1}{2}\begin{bmatrix}
-3 & -3 & 0 & 1 & 0\\
-1 & -1 & -2 & -1 & 0\\
-1 & -1 & 0 & 1 & -2
\end{bmatrix} \tag{2.44}$$

2.3.6 The Projection Operators $\mathcal{C}$ and $\mathcal{O}$

Returning to (2.42), one observes that the matrix equation satisfies the conditions of a projection operator. In other words, $\mathcal{C} = \mathsf{G}\,\mathsf{Q}$ can be observed as a projection operator $\mathcal{C} : \mathbf{P}^p(\triangle) \rightarrow \nabla P^{p+1}(\triangle) = \mathbf{K}^p$. Using $\mathcal{C}$, any polynomial in $\mathbf{P}^p(\triangle)$ can be projected into $\mathbf{K}^p$. Now, one can see that $\mathcal{O}$ defined in (2.45), satisfies $\mathcal{C}^T\mathcal{O} = 0$. In other words, any vector from the range space of $\mathcal{O}$ is orthogonal to any other vector from the range space of $\mathcal{C}$. Note that the orthogonality is with respect to the inner product of the associated Degree(s) of Freedom (DoF) row vectors E. Hence, considering the finite dimensionality of $\mathbf{P}^p(\triangle)$, an orthogonal decomposition of the form (2.5) is obtained. Moreover, as formulated in (2.45), $\mathcal{O}$ is a projection operator $\mathcal{O} : \mathbf{P}^p(\triangle) \rightarrow [\nabla P^{p+1}(\triangle)]^{\perp} = \mathbf{K}^{p\perp}$.

$$\mathcal{O} = \left[\begin{array}{c|c} 0_{tt} & -\mathsf{Q}_t^T\mathsf{G}_c^T \\ \hline 0_{ct} & \mathsf{I}_{cc} \end{array}\right]_{d\times d} = \mathsf{I}_{d\times d} - \mathcal{C}^T,\ d \triangleq \dim \mathbf{P}^p(\triangle) \tag{2.45}$$

2.4 Global Assemblage of the Constraints

This section explains the method that is used for assembling the constraint equations from the metric independent matrix C of section 2.3.5. First, it is needed to have a general idea about the dimensionality of the discrete (polynomial) spaces involved.

2.4.1 Dimensionality

Assuming a fixed triangulation $\mathcal{T}$ over the problem domain $\mathbf{\Omega}$, let E and N respectively denote the number of elements and vertex nodes in $\mathcal{T}$. Also, assume that a spanning tree over the *primal* graph $\mathcal{G}$ of $\mathcal{T}$ is given and that the numbers of edges, tree edges and cotree edges are respectively denoted by e, e_t and e_c. Using *Euler* identity, one concludes that $e_t = N - 1$ and $e_c = E$. Considering the number of Degree(s) of Freedom (DoF) on individual edges and faces, (2.46) yields the dimension of a $(p+1)$-order $H^1(\mathbf{\Omega})$ polynomial space over $\mathcal{T}$. Similarly, the dimension

of a p-order $H(\nabla\wedge,\boldsymbol{\Omega})$ polynomial expansion is given by (2.47).

$$\dim \boldsymbol{\Pi}^{p+1}(\boldsymbol{\Omega}) = ep + Ep(p-1)/2! + N = p(e + e_c(p-1)/2!) + e_t + 1 \tag{2.46}$$

$$\dim \mathbf{V}^p(\boldsymbol{\Omega}) = e(p+1) + E\,p(p+1) = (p+1)(e + e_c\,p) \tag{2.47}$$

Since $\ker(\nabla, \boldsymbol{\Pi}^{p+1})$ is the one dimensional polynomial space of constant scalar fields, using rank nullity theorem and (2.6), the dimension of $\nabla\boldsymbol{\Pi}^{p+1}$ is obtained as (2.48). Following the discussion of section 2.3.6, $\dim \mathbf{K}^{p\perp}$ can be found using (2.49).

$$\dim \mathbf{K}^p = \dim \nabla\boldsymbol{\Pi}^{p+1}(\boldsymbol{\Omega}) = \dim \boldsymbol{\Pi}^{p+1}(\boldsymbol{\Omega}) - 1 = p(e + e_c(p-1)/2!) + e_t \tag{2.48}$$

$$\dim \mathbf{K}^{p\perp} = \dim \mathbf{V}^p - \dim \mathbf{K}^p = e_c(p+1)(p+2)/2! \tag{2.49}$$

2.4.2 Global Assemblage

In this section, we shall attempt to extend the (local) tree/cotree splitting of Degree(s) of Freedom (DoF) (see section 2.3.3) into the global level. At the same time, we shall propose a method for the construction of the constraint equations (analogous to those of section 2.3.5) at the global Degree(s) of Freedom (DoF) level. Turning back to (2.49), it is understood that for every element (and every cotree edge) in $\mathcal{T}$, there must be $(p+1)(p+2)/2!$ Degree(s) of Freedom (DoF) in $\mathbf{K}^{p\perp}$. In other words, for every element in $\mathcal{T}$, $(p+1)(p+2)/2!$ linearly independent equations must be specified so that $\mathbf{K}^p$ (or the orthogonal *complement* of $\mathbf{K}^{p\perp}$ in $\mathbf{V}^p$) is entirely determined in $\mathbf{V}^p$. A global version of the gradient/path integration approach can still be deployed for the

development of the required constraint equations. This has been adopted in [58], though for a slightly different purpose and limited to the least polynomial order. Yet, from complexity (and accuracy) point of view, it is preferred if the desired constraint equations can be extracted from the single-element constraints matrix C of (2.43). Nevertheless, while in $\mathbf{K}^p$, the constraint equations of (2.43) should still be valid locally over individual elements. Considering that $e_c = E$, a proposed subset of Degree(s) of Freedom (DoF) is acceptable as the set of (global) cotree Degree(s) of Freedom (DoF) if it fulfills the following conditions:

1. There should be exactly $e_c(p+1)(p+2)/2!$ Degree(s) of Freedom (DoF) in the set.

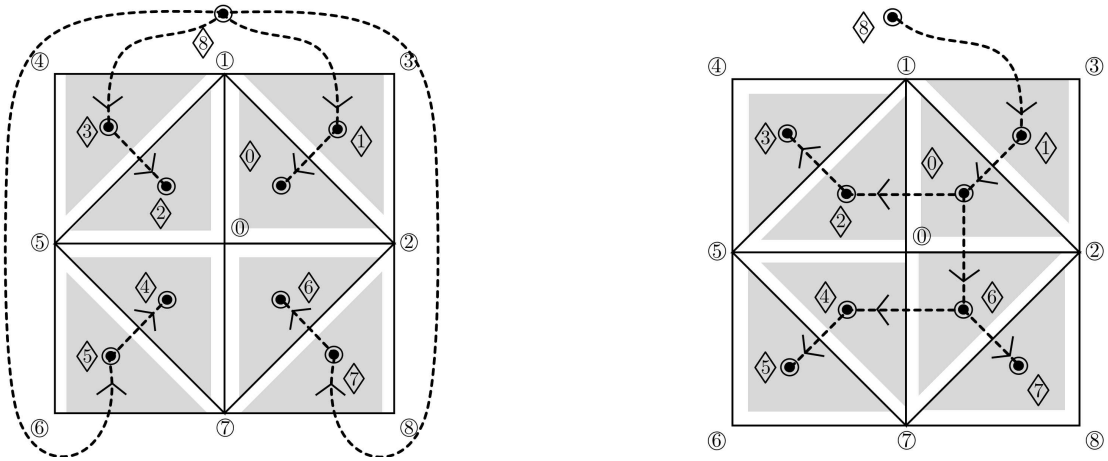
2. For every element, $(p+1)(p+2)/2!$ appropriate equations (constraints) must be specified. Each of these constraints must express one and only one of the proposed cotree Degree(s) of Freedom (DoF) in terms of Degree(s) of Freedom (DoF) that solely belong to the complement of the proposed set.

3. For every element, the proposed $(p+1)(p+2)/2!$ constraints should be derived from individual element constraints matrix C.

4. For every element, the proposed $(p+1)(p+2)/2!$ constraint equations must form a linearly independent set.

Using figure 2.5b as an example, the approach that fulfills the above requirements is proposed. In the figure, a spanning tree is constructed over the dual grid such that the tree is rooted at exteriority. This results in having a primal spanning tree dually defined on the primal grid.

Now, assume that individual element basis is settled in such a way that the dual tree branch which enters each element points to the local node ① of the same element. With this assumption, the global 'cotree Degree(s) of Freedom (DoF)' shall be the union of individual elements' local cotree Degree(s) of Freedom (DoF) described in section 2.3.3. Further, the set of 'global tree' Degree(s) of Freedom (DoF) is the set complement to the global cotree Degree(s) of Freedom



(a) The relatively wide tree on the FE mesh.

(b) The relatively tall tree on the FE mesh.

Figure 2.5: Two choices for a spanning tree on the dual grid. Nodes numbers and element numbers are indicated by small ◯ and ♢ signs respectively. The gray sub-triangles symbolically refer to the global cotree DoF residing in each element.

(DoF). Thus, item 1 is satisfied. Furthermore, on elements that fall on the ends of dual tree branches (elements $\{3,5,7\}$ in figure 2.5b's example), the local constraints matrix C gives $(p+1)(p+2)/2!$ equations that satisfy item 2, item 3 and item 4. This is true since for such elements the global tree Degree(s) of Freedom (DoF) coincide with (individual) local tree Degree(s) of Freedom (DoF). Now, if one uses the local C to achieve the required $(p+1)(p+2)/2!$ for elements that are located one 'level' before the branch ends, the resulting equations would express the element cotree Degree(s) of Freedom (DoF) in terms of a mixture of global tree and cotree Degree(s) of Freedom (DoF). This happens because the local tree Degree(s) of Freedom (DoF) falling at least on one of the two local tree edges already belong to the global cotree Degree(s) of Freedom (DoF). For example, in figure 2.5b, the Degree(s) of Freedom (DoF) falling on edges { ②①, ①⑤, ⑤⑦, ⑦② } of elements $\{0,2,4,6\}$ belong to the cotree Degree(s) of Freedom (DoF). But, these cotree Degree(s) of Freedom (DoF) have already been expressed in terms of global tree Degree(s) of Freedom (DoF) when elements on the branch ends were being processed. This means that such global cotree Degree(s) of Freedom (DoF) can be re-expressed using the constraints equations developed in the previous stage or 'level'. The process can be continued, level by level, until all elements are processed and all of the item 1 through item 4 are fulfilled.

Technically speaking, along this process, it is necessary to have all of the children of an individual element processed before getting involved with the element itself. Hence, the process begins from the branch ends and terminates when the root is reached. The algorithm can be implemented in either a recursive or level based approach. However, the level based method, as implemented in this work, is much more effective. This is further discussed in section 2.7.

Now, having the global Degree(s) of Freedom (DoF) partitioned into tree and cotree Degree(s) of Freedom (DoF), in analogy to the single element case discussed in section 2.3.5 and section 2.3.6, one arrives at a global constraints matrix satisfying (2.50).

$$\mathsf{E}_c = \mathsf{C}\,\mathsf{E}_t \tag{2.50}$$

Consequently, the global $\mathcal{C}$ projector is derived in (2.51). Nevertheless, global $\mathcal{O}$ projector still obeys (2.45).

$$\left[\begin{array}{c}\mathsf{E}_t\\\hline\mathsf{E}_c\end{array}\right] = \mathcal{C}\left[\begin{array}{c}\mathsf{E}_t\\\hline\mathsf{E}_c\end{array}\right],\ \mathcal{C} \triangleq \left[\begin{array}{c|c}\mathsf{I}_{tt} & \mathsf{0}_{tc}\\\hline\mathsf{C} & \mathsf{0}_{cc}\end{array}\right] \tag{2.51}$$

Finally, it is worth emphasizing that the choice of the dual tree is non-unique. The mentioned tree and cotree Degree(s) of Freedom (DoF), and consequently the resulting constraint equations, are non-unique. See figure 2.5a for another example of the dual grid compared to that of figure 2.5b.

2.5 Relaxing C for Physical Gradient Modes

In presence of multiple disjoint equipotential surfaces or PECs, the true solutions of Maxwell equations cease to exclude all gradient fields. It is known [68] that the number of such gradient modes is one less than the total number of the disjoint PEC surfaces. Under these conditions, the space of unwanted (nonphysical) gradient fields shrinks; the gradient fields in which the PEC

surfaces obtain non-equal scalar potentials [68] are excluded from the unwanted solution space. More precisely speaking, the unwanted solution space is the space of gradient fields in which all PEC surfaces have equal potentials.

As argued in the previous section, such physical '*direct current (DC)* ' solutions belong to $\mathbf{K}^p$. If the FEM solution is enforced to completely exclude $\mathbf{K}^p$ components, all gradient modes will be eliminated regardless of their physical or non-physical nature. Hence, $\mathbf{K}^p$ must be appropriately shrunk (restricted) such that the physical gradient solutions are excluded from $\mathbf{K}^p$ and included in the space of desirable solutions. This must be realized by properly modifying the projection operators or the C matrix. Since C represents a set of constraints, reducing the number of constraints enforced by C is here regarded as 'relaxing'. Note that this is an exact reduction of the rank of C and no approximation is involved in this 'relaxation' process.

Now, assume that m disjoint PEC surfaces exist in the problem. The problem possesses (precisely) $m-1$ physical gradient solutions [68]. Assume ϕ_1, ϕ_2, ..., ϕ_m to denote the scalar Degree(s) of Freedom (DoF) associated with m nodes of the FEM grid such that each of the nodes belongs to one and only one of the PEC surfaces and coincides with a triangle vertex. Denote the subspace of gradient fields in which (2.52) is satisfied by $\mathbf{K}^p_N \subset \mathbf{K}^p$.

$$\phi_1 = \phi_2 = \ldots = \phi_m \tag{2.52}$$

It is clear that $\dim \mathbf{K}^p_N = \dim \mathbf{K}^p - (m-1)$. Moreover, denote the orthogonal complement of $\mathbf{K}^p_N$ in $\mathbf{K}^p$ by $\mathbf{K}^p_P = \mathbf{K}^{p\perp}_N$. It is immediately concluded that $\dim \mathbf{K}^p_P = (m-1)$. Clearly, before enforcing the PEC conditions, different nodes on individual PEC surfaces are allowed to possess non-equal potentials. Indeed, $\mathbf{K}^p_N$ and $\mathbf{K}^p_P$ can appropriately model the spaces of nonphysical and physical gradient solutions only after the PEC conditions are completely enforced.

The objective here is to modify the constraint equations such that they only model the non-physical subspace $\mathbf{K}^p_N$. Taking ϕ_m as a reference, one can find $m-1$ paths on the primal FEM

tree that connect the nodes associated with $\phi_1, \phi_2, \ldots, \phi_{m-1}$ to the node associated with ϕ_m. Note that the mentioned primal tree must be dual to the dual tree of choice used in the previous section(s). If (2.52) is to be satisfied by the elements of $\mathbf{K}_N^p$, the electric field integral on each of the mentioned paths must be zero. Translating this physical condition into algebraic equations yields a set of $m-1$ linear equation entirely in terms of the 'tree' Degree(s) of Freedom (DoF):

$$\begin{cases} u_{11}\mathsf{E}_{t1} + u_{12}\mathsf{E}_{t2} + \ldots + u_{1\,\dim\mathbf{K}^p}\mathsf{E}_{t\,\dim\mathbf{K}^p} = 0 \\ \vdots \\ u_{m-1,1}\mathsf{E}_{t1} + u_{m-1,2}\mathsf{E}_{t2} + \ldots + u_{m-1,\dim\mathbf{K}^p}\mathsf{E}_{t\,\dim\mathbf{K}^p} = 0 \end{cases} \tag{2.53}$$

In (2.53), each equation allows for one of the tree Degree(s) of Freedom (DoF) to be expressed as a linear combination of one or more other tree Degree(s) of Freedom (DoF). Such equations allow for re-expressing an $(m-1)$ member subset of E_t's Degree(s) of Freedom (DoF) in terms of its other Degree(s) of Freedom (DoF). This leads to a decomposition of E_t's Degree(s) of Freedom (DoF) into say E_u and E_d with a block matrix X_{du} relating E_u and E_d in (2.54) where E_u is an $(m-1)$ dimensional block vector.

$$\mathsf{E}_u = \mathsf{X}_{ud}\mathsf{E}_d \tag{2.54}$$

With this new decomposition, (2.50) and (2.51) will be modified to (2.55) and (2.56)).

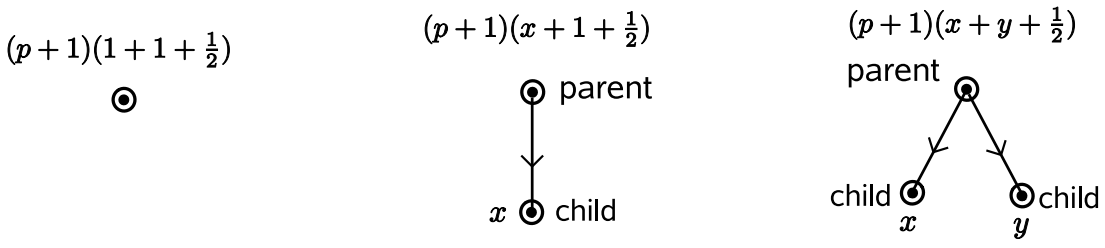
$$\begin{cases} \mathsf{E}_d = \mathsf{I}_{dd}\mathsf{E}_d \\ \mathsf{E}_u = \mathsf{X}_{ud}\mathsf{E}_d \\ \mathsf{E}_c = \mathsf{C}_{cd}\mathsf{E}_d + \mathsf{C}_{cu}\mathsf{E}_u \end{cases} \tag{2.55}$$

$$\left[\begin{array}{c} \mathsf{E}_d \\ \hline \mathsf{E}_u \\ \hline \mathsf{E}_c \end{array}\right]_{\mathbf{K}_N^p} = \mathcal{C} \left[\begin{array}{c} \mathsf{E}_d \\ \hline \mathsf{E}_u \\ \hline \mathsf{E}_c \end{array}\right]_{\mathbf{V}^p}, \; \mathcal{C} \triangleq \left[\begin{array}{c|c|c} \mathsf{I}_{dd} & 0_{du} & 0_{dc} \\ \hline \mathsf{X}_{ud} & 0_{uu} & 0_{uc} \\ \hline \mathsf{C}_{cd} + C_{cu}X_{ud} & 0_{cu} & 0_{cc} \end{array}\right] \tag{2.56}$$

It is worth emphasizing here that we are developing the constraints expressing the functional space of undesired gradient FE solutions $\mathbf{K}_N^p$. Still, the global projector $\mathcal{O}$ can be obtained using (2.45). The essential difference, however, is a reduction of the effective rank of C by $m-1$. This reduction, allows for $m-1$ extra dimensions to attach to the space of 'desired' FEM functions. Take note that the modified projections now have different range spaces: $\mathcal{O}:\mathbf{V}^p\longrightarrow\mathbf{K}_N^{p\perp}\oplus\mathbf{K}^{p\perp}$, $\mathcal{C}:\mathbf{V}^p\longrightarrow\mathbf{K}_N^p$. These have been denoted with $\mathbf{V}^p$ and $\mathbf{K}_N^p$ subscripts in (2.56).

2.6 Sparsity

In practice, the global projection operator $\mathcal{O}$ is built out of the global constraints matrix C. This operator is used to project the RHS of the final matrix problem before Krylov iterations start. Thus, it imposes the extra computational burden of one matrix-vector multiplication per matrix solution. This situation becomes more critical when eigenvalue problems are to be dealt with; The so called shift and invert approach requires repeated solution of the matrix problem. In any of the mentioned situations, efficient assemblage of the sparse matrix is of paramount importance. The global constraints matrix C, inherits $(p+1)(p+2)/2!$ equations(rows) from each individual element. Following the assemblage process of section 2.4.2, the number of columns involved in each element's constraint equations depends on the number of columns involved in the constraint equations of its downstream elements (children) on the dual tree. In this regard, three different situations arise:



(a) no children. (b) 1 child. (c) two children.

Figure 2.6: The sparsity (the number of nonzero columns in each of the constraint equations) of an individual element's constraint equations depends on the sparsity of constraint equations in its direct descendant(s).

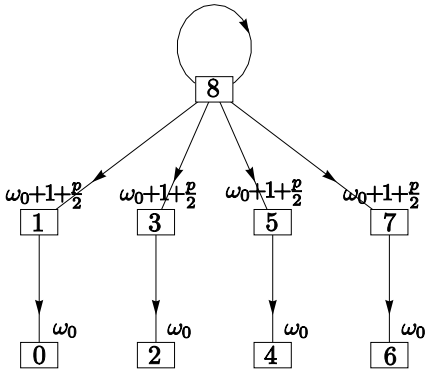
1. The element has no children: $(p+1)(1+1+p/2)$ columns are involved in each equation.

2. The element has a child: $(p+1)(x+1+p/2)$ columns are involved in each equation where $(p+1)$ x is the number of columns involved for the child.

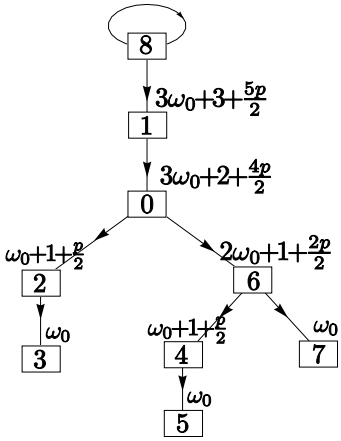
3. The element has two children: $(p+1)(x+y+p/2)$ columns are involved in each equation where $(p+1)$ x and $(p+1)$ y are the number of columns involved for the two children.

These three situations are visualized in figure 2.6. Note that the number of nonzero columns in all three cases is a multiple of $(p+1)$. Using the diagrams of figure 2.6 it can be observed that the sparsity of the constraints matrix C is directly affected by the choice of the dual tree. Figure 2.7 demonstrates the effect of the choice of the dual tree on the number of nonzero columns involved in each of the $(p+1)(p+2)/2!$ constraint equations expressed by individual elements. The (dual) trees depicted in figure 2.7 are identical to those of figure 2.5a and figure 2.5b. As indicated in figure 2.6a, for elements that reside at the branch ends (of the dual tree), the number of nonzero columns involved in each of the constraint equations equals $(p+1)(1+1+\frac{1}{2})$. Factoring the common $(p+1)$ term, let's define $\omega_0 = (1+1+\frac{1}{2})$. With this definition, the values expressed at individual nodes in figure 2.7 represent the number of nonzero columns in each of the equations expressed by individual nodes (actual elements) of the tree. Note that for the sake of compactness a common $(p+1)$ term is factored from all of the terms depicted in figure 2.7. Considering that $(p+1)(p+2)/2$ equations are expressed at each node of the diagram and summing up the depicted values, it is revealed that the wide tree of figure 2.5a (and figure 2.7a) constitutes to a global C matrix with $\frac{1}{2}(p+1)^2(p+2)(13\omega_0+8+13\frac{p}{2})$ nonzero entries. This is while the tall tree of figure 2.5b (and figure 2.7b) gives rise to $\frac{1}{2}(p+1)^2(p+2)(8\omega_0+4+4\frac{p}{2})$ nonzero entries for C. Factoring the common terms and simplifying the sums one arrives at $13\omega_0+8+13\frac{p}{2} = 34+13p$ nonzero entries for figure 2.5b versus $8\omega_0+4+4\frac{p}{2} = 20+6p$ nonzero entries for figure 2.5a. In essence, the tall tree has more nonzero entries as compared to the wide tree. However, the asymptotic ratio between different tree choices can be much larger for meshes comprising of

large number of elements. Consequently, when the global number of Degree(s) of Freedom (DoF) grows, the necessity for an optimal choice of the dual tree cannot be ignored. Our experiments with various sample meshes reveal that the number of non-zeros in C can easily and by far exceed the non-zeros of the LHS and RHS FEM matrices.



(a) The (wide) tree of figure 2.5a



(b) The (tall) tree of figure 2.5b

Figure 2.7: Number of columns involved in each of the $(p+1)(p+2)/2!$ constraint equations expressed by individual elements (or cotree edges). Here ω_0 is defined as $\omega_0 \triangleq (1+1+p/2)$. Note that a $(p+1)$ term is factored out and the true number of columns must be obtained by multiplying the numbers in the diagram by $(p+1)$.

2.7 Parallelization

This section is attempting to document the approach that has been actually deployed for a parallel assemblage of the constraints matrix on a conventional[3] dual-core *central processing unit (CPU)* system. Future implementations of the implemented algorithm are hence expected to be improved. The focus of this section remains on the fact that the constraints matrix is parallelizable with the aid of splitting the dual tree into the so called subtree components. Parallel assemblage of FEM matrices is becoming common practice due to the availability of low-cost parallel hardware and open source parallel matrix processing software like *Epetra* [4] [50]. A parallel matrix is basically a sparse data structure in which certain rows are assigned to a particular process. The assignment of

[3] Intel Pentium D

[4] http://trilinos.sandia.gov/

matrix rows into processes is often regarded as the *map* (more precisely the 'the row map') of the parallel matrix. Inappropriate choice of the map leads to unnecessary *off-process* communications and imbalanced system loading. For the constraints matrix of our concern, however, the map must be constructed with respect to the dual tree. During the assemblage of parallel matrices, it is best to avoid reading from the incompletely filled matrix. This is true because of two main reasons:

- When the entries are still not assembled (still hanging somewhere in a buffer), reading may be fatal. The entries may yet be inaccessible specially if they belong to other processes.
- Even assuming that the reading can be correctly handled, it would not impose the usual costs of accessing the assembled matrix entries.

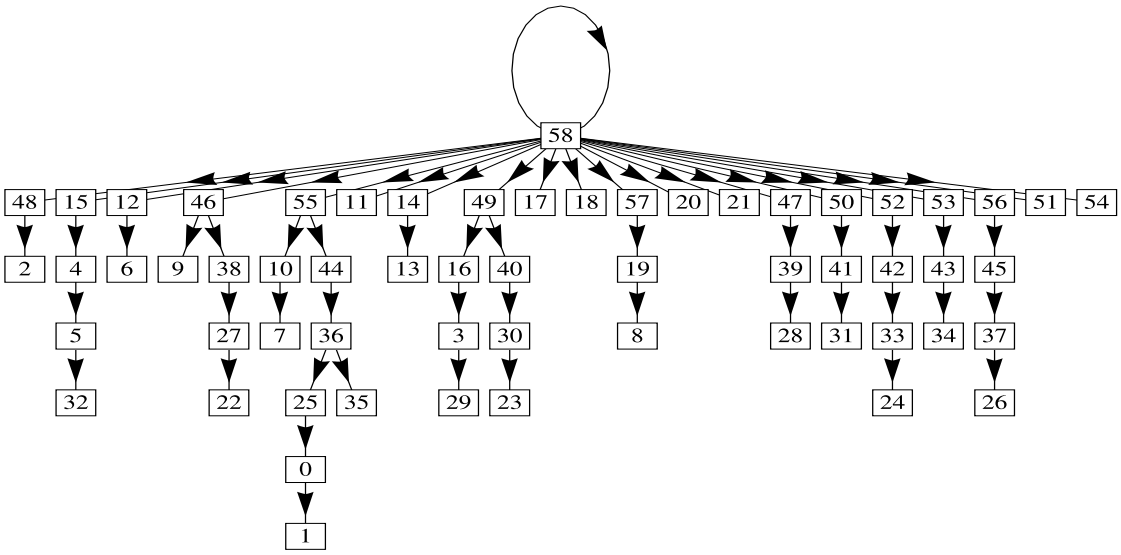


Figure 2.8: The dual tree associated to the primal FEM mesh of figure 2.9.

It is possible to assemble the matrix using a recursive function that starts the call from the top of the dual tree and terminates while reaching back from the bottom. However, it wold be extremely inefficient and memory demanding due to memory allocations in the pending levels and the recursive nature of the algorithm itself. Yet, a more efficient way of constructing the constraints matrix is to use the critical fact that When the elements of a certain layer are being precessed, only the entries (rows) that were constructed during the previous (immediately one

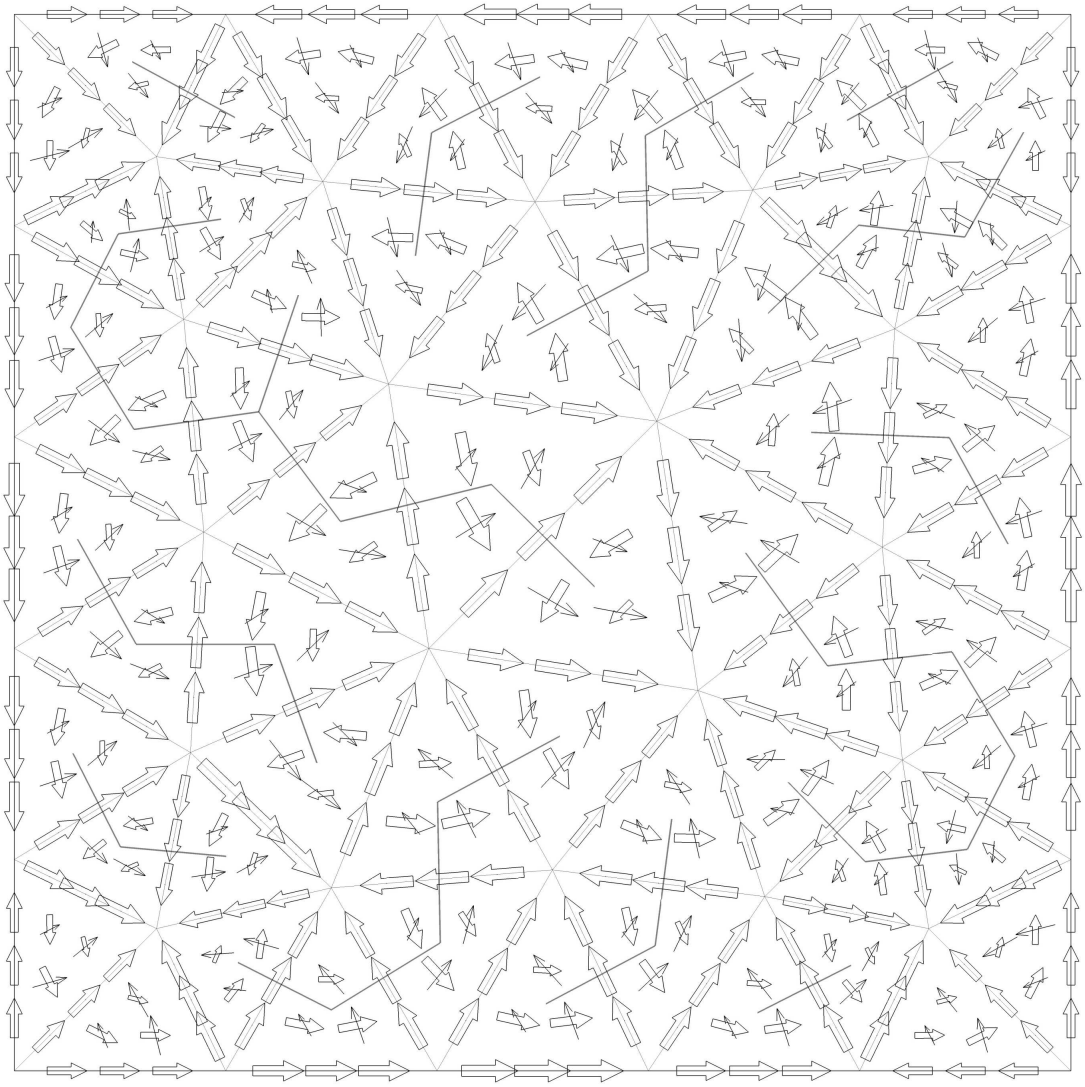


Figure 2.9: A FEM mesh and the associated dual tree. Note the lines that connect the elements into exteriority are not plotted in this computer visualization. Vector DoF are also symbolically depicted.

layer below) layer are needed. Thus the matrix can be actually assembled using the following algorithm:

1. Start from the bottom layer of the tree (see for example figures 2.8, 2.9 and 2.10).

2. Based on the critical fact stated in above, the calculated matrix rows must be maintained (except from what is being assembled into the global matrix) in a buffer till the assemblage for the next layer is completed. This process continues until the top (root) element in the dual tree is reached. The root element is a virtual element which represents the exteriority

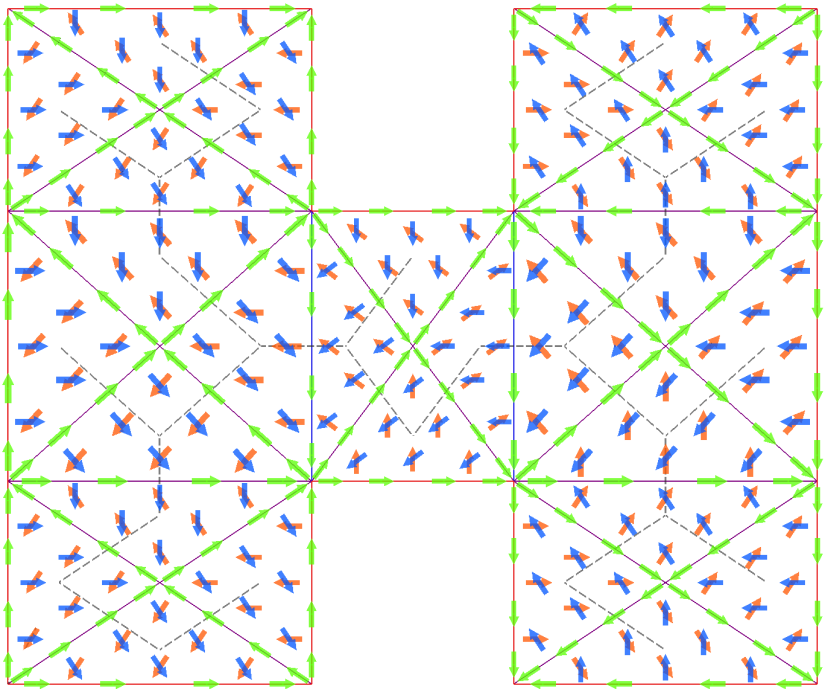
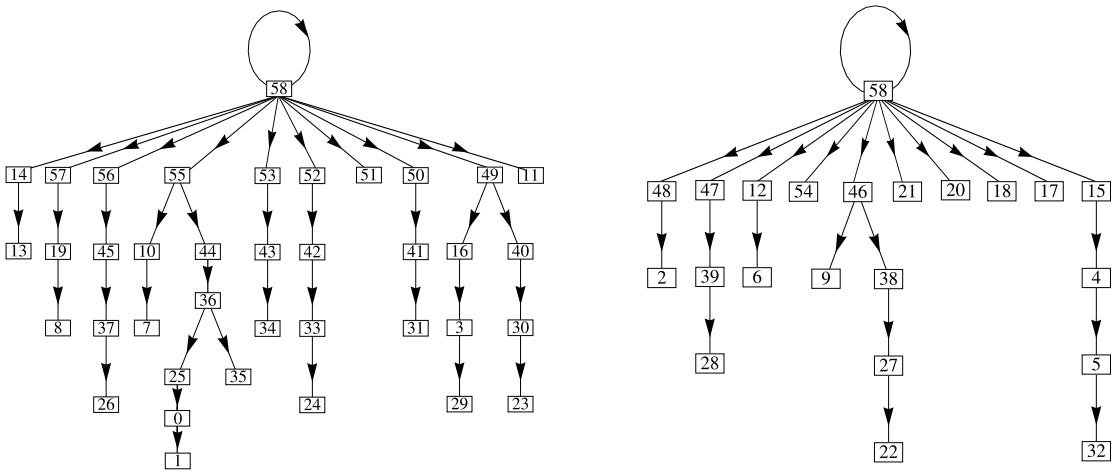


Figure 2.10: Another FEM mesh and the associated dual tree. Note the lines that connect the elements into exteriority are not plotted in this computer visualization. Vector DoF are also symbolically depicted.

to the primal mesh.

From parallelization point of view, the dual tree can be decomposed into independent *subtrees* as exemplified in figure 2.11. Each process can then be assigned one or more subtrees. In this way, every process precisely possesses the matrix rows corresponding to the cotree Degree(s) of Freedom (DoF) defined by its associated subtree(s). Hence, any off-process communication is avoided. Furthermore, for the purpose of efficient memory management and symmetric load balancing, the main tree can be split into many subtrees reducing the maximum width of the subtrees. This reduces the size of the buffer that is used for handing matrix rows between consecutive levels. Furthermore, a symmetric load balance can be achieved by means of properly dispatching the subtrees among available processes. Even in case of a serial implementation, subtree based assemblage leads to substantial memory savings. This is because the size of the buffer needed for storing the data from the previous layer is proportional to the (maximum) width of the tree (subtrees) being processed. Our primitive experiments with the approach indicate that it leads



(a) Part1. (b) Part2.

Figure 2.11: The main dual tree is decomposed into two independent subtrees, each of which correspond to an independent set of lines of the constraints matrix C.

considerable savings in the required computational resources.

2.8 Feasibility of Extension to 3D

Due to the natural properties of simplices, a three dimensional (in general n-dimensional) extension of the proposed T/C decomposition is completely feasible. In this regard, two key features must be identified for the $3D$ analogue: the dual graph and the rotational symmetry of the basis discussed in section 2.3.3.

2.8.1 The Structure of the Dual Tree

The analogue definition of the dual graph for the $3D$ case is straightforward:

1. Every element (and the exteriority domain) becomes a node in the dual graph. The node is then regarded as 'dual to' the associated element.

2. Every face (of tetrahedral elements) turns into an edge connecting the nodes that are dual to its adjacent elements. The edge will be regarded as the dual of the associated face

Since a d-simplex has $\binom{d+1}{d} = (d+1)$ hyperfaces and since each of the hyperfaces are in one-to-one correspondence with dual tree branches it can be concluded that in the d-dimensional case, the dual tree will at most have d branches at each junction (except at the exteriority junction). Similarly, if the approach is extended to hypercubic elements (i.e. quadrilaterals in *two dimensional/dimensions (2D)* and bricks in $3D$) the tree will at most have $(2\ d-1)$ (one less than the total number of each element's faces) branches at each junction (except at the exteriority junction).

2.8.2 Tree/Cotree DoF at Element Level

In $3D$, Nedelec's zero order $H(\nabla\wedge)$ conforming basis comprises of $\binom{d+1}{2} = \binom{2+1}{2} = 6$ components: $\{\Omega_{12}, \Omega_{13}, \Omega_{14}, \Omega_{23}, \Omega_{24}, \Omega_{34}\}$ where each of the Ω_{uv}s follow the definition in (2.7). The interpolatory basis is constructed by multiplying appropriate Lagrangian polynomials with appropriate choices of Ω_{uv} [12]. Since, we are concerned with the rotational symmetry of the spatial orientation of the basis, we will focus on the following question: which Ω_{uv} must be used for the Degree(s) of Freedom (DoF) associated with particular edge, face and internal nodes so that the desired rotational symmetry of basis can be achieved? To answer this question, we will take advantage of the following lemma. Note that a d-dimensional simplex is identified by its $(d+1)$ vertices denoted by ⓘ, $1 \leq i \leq d+1$.

Lemma 2.8.1 *For every d-dimensional simplex, all $\Omega_{uv}, u, v \neq o$ are perpendicular to any line passing through vertex ⓞ.*

Proof We shall provide the proof for $o = 1$. Other cases are justified by properly permuting the vertices of the simplex. Assume that the *Cartesian* coordinates and unit directional vectors in $\mathbb{R}^d$ are denoted by x_i and $\hat{x}_i$ respectively. Consider that ① is located at the origin of the Cartesian coordinates and that the other vertices reside at ⓡ : $x_i = \delta_{r-1,i}, 2 \leq i \leq d+1$. Denote the tangent to the line originating from vertex ① and passing through any point p : $\sum_{i=1}^{d} \xi_{i+1}\hat{x}_i$

by τ. τ would then be parallel to $\sum_{i=1}^{d}\xi_{i+1}\hat{x}_i$ in $\mathbb{R}^d$. With the same assumptions, $\nabla\xi_r = \hat{x}_{r-1}, 2 \leq r \leq d+1$. Thus $\Omega_{uv}, u, v \neq o$ will be equal to $\xi_u\hat{x}_{v-1} - \xi_v\hat{x}_{u-1}$ in $\mathbb{R}^d$. Consequently $\tau.\Omega_{uv} = 0, u, v \neq o$.

As a result of lemma 2.8.1, one can state that polynomial basis components aligned along $\Omega_{uv}, u, v \neq o$ are perpendicular to the line originating from vertex ⓪. In other words, the path integration (on a trajectory that starts from vertex ⓪) of such basis components will be zero. This property was exploited in section 2.3.3 where it was regarded as the rotational symmetry of the basis around vertex ⓪.

Now we return to our main question: which Ω_{uv} must be used for the Degree(s) of Freedom (DoF) associated with particular edge, face and internal nodes so that the desired rotational symmetry of the basis can be achieved? Based on the observation made in lemma 2.8.1 the answer to this question is brought forward in the following enumeration. Note that the design is intended for building the rotational symmetry around node ① of the $3D$ simplex.

1. For each edge (uv), the associated Degree(s) of Freedom (DoF) (one at a node) must be constructed using Ω_{uv}

2. For each face (uvw) where ① $\in \{u, v, w\}$, the associated Degree(s) of Freedom (DoF) must be constructed using $\Omega_{u'v'}$s (two Degree(s) of Freedom (DoF) at a node) where $u'v'$ can be any nondegenerate combination of the $\{u, v, w\}$ subject to the constraint that exactly one occurrence of $u'v'$ includes ①.

3. For each face (uvw) where ① $\notin \{u, v, w\}$, the associated Degree(s) of Freedom (DoF) must be constructed using $\Omega_{u'v'}$s (two Degree(s) of Freedom (DoF) at a node) where $u'v'$ can be any nondegenerate combination of the $\{u, v, w\}$.

4. For internal nodes, the associated Degree(s) of Freedom (DoF) must be constructed using $\Omega_{u'v'}$s (three Degree(s) of Freedom (DoF) at a node) where $u'v'$ can be any nondegenerate

combination of the $\{①, ②, ③, ④\}$ subject to the constraint that exactly one occurrence of $u'v'$ includes ①.

If the abovementioned conditions are met, (in analogy to section 2.3.3) one can observe that in the discrete path integration operator, the columns associated with the following Degree(s) of Freedom (DoF) will be entirely zero.

1. All Degree(s) of Freedom (DoF) on the single face (234).
2. Two out of the three Degree(s) of Freedom (DoF) associated with internal nodes.
3. One out of the two Degree(s) of Freedom (DoF) associated with the faces $(1uv)$.

The observation can be justified using lemma 2.8.1 and considering the spatial orientation of various basis components as designed in the enumeration above. In analogy to section 2.3.3, the Degree(s) of Freedom (DoF) listed above will form the cotree Degree(s) of Freedom (DoF). Excluding the cotree Degree(s) of Freedom (DoF), there will remain one Degree(s) of Freedom (DoF) for every node that is neither on face (234) nor on vertex ①. Thus for a p-order $H(\nabla\wedge)$ basis, there would be $\binom{p+1+d}{d}-1$ such Degree(s) of Freedom (DoF) regarded as the 'tree' Degree(s) of Freedom (DoF). This is exactly one less than the dimension of a $(p+1)$-order H^1 basis over the same simplex. The equality is by no means coincidental and is the direct consequence of (2.6). Figure 2.12 gives examples of the interpolation nodes for construction of the mentioned $(p+1)$-order H^1 and p-order $H(\nabla\wedge)$ bases over the unit isometric tetrahedral (for $p=2$).

2.9 Numerical Examples

Krylov subspace solvers were used for the results presented in this section. For the eigenproblems, the *Block Krylov Schur (BKS)* solver of the *Anasazi* package from the Trilinos project [50] was used. Since the smallest eigenvalues are practically desired, a shift and invert approach

is implemented which equivalently seeks the largest eigenvalues of $\frac{1}{\lambda-\sigma}\mathsf{B}\,\mathsf{x} = (\mathsf{A}-\sigma\mathsf{B})\,\mathsf{x}$ instead of seeking the smallest eigenvalues of $(\mathsf{A}-\sigma\mathsf{B})\mathsf{x} = (\lambda-\sigma)\mathsf{B}\mathsf{x}$. For the boundary value problem of section 2.9.3, various solvers including *preconditioned conjugated gradient (PCG)* and *preconditioned generalized minimal residual (PGMRES)* were examined using the *AztecOO* package which is again obtained as a sub-package of the Trilinos project [50].

The method by which $\mathcal{O}$ is incorporated for eliminating the unwanted modes is straightforward. In either of the aforementioned cases, the user is required to provide the means for solving (2.57) or (2.58) for x.

$$(\mathsf{A}-\sigma\mathsf{B})\,\mathsf{x} = \mathsf{B}\,\mathsf{y} \tag{2.57}$$

$$\mathsf{A}\,\mathsf{x} = \mathsf{y} \tag{2.58}$$

Using *Galerkin* method, the FEM matrices can be formulated as (2.59) and (2.60) [58]. Here, $\mathbf{W}_i$ represent the FEM global vector basis while μ_r and ϵ_r refer to relative magnetic permeability and electric permittivity tensors as element-wise constant functions of space.

$$\mathsf{A}_{ij} = \int_{\mathbf{\Omega}} (\nabla\wedge\mathbf{W}_i)\cdot(\mu_r^{-1}\nabla\wedge\mathbf{W}_j)\mathrm{d}\mathbf{\Omega} \tag{2.59}$$

$$\mathsf{B}_{ij} = \int_{\mathbf{\Omega}} (\mathbf{W}_i)\cdot(\epsilon_r\mathbf{W}_j)\mathrm{d}\mathbf{\Omega} \tag{2.60}$$

From (2.59), it is concluded that $\ker(\nabla\wedge,\mathbf{V}^p) \cong \mathrm{null}(\mathsf{A})$. In other words, the row vector representation of $\ker(\nabla\wedge,\mathbf{V}^p)$ (as a subspace of $\mathbf{V}^p$) forms the null-space of A. When solving (2.57) in the presence of a proper $\mathrm{null}(\mathsf{A})$, it is necessary to confine the RHS $\mathsf{B}\,\mathsf{y}$ to the range-space of A. Similarly, when solving (2.58), y must be confined to the range-space of A. Now, whatever the solver process is going to be, say PCG or PGMRES, this can be achieved by pre-

multiplying y or B y by the projection operator $\mathcal{O}$. The only difference between the two cases is that, for the eigensolver the mentioned 'linear solve' operation is performed repeatedly upon the requests from the reverse communication interface of the BKS solver while for the boundary value problem this linear solve is performed only once.

2.9.1 Hollow Waveguide

The first example presented here involves the convergence study of a rectangular hollow waveguide with its width and height dimensions $a = 3$ and $b = 2$ respectively. In the presence of natural boundary conditions on the periphery, the eigenvalues can be analytically found from formula (2.61) which can be obtained from [75].

$$\kappa_{mn}^2 = \pi^2(\frac{m^2}{a^2} + \frac{n^2}{b^2}); m, n \geq 1 \tag{2.61}$$

Hence, the first ten analytical values for κ^2 are $\{\frac{13}{36}, \frac{25}{36}, \frac{10}{9}, \frac{5}{4}, \frac{13}{9}, \frac{2}{1}, \frac{73}{36}, \frac{85}{36}, \frac{97}{36}, \frac{25}{9}\}\pi^2$. As expected from the discussion above, the introduction of $\mathcal{O}$ operator into the matrix solver resulted in complete elimination of spurious modes. Hence, the dominant (smallest) eigenvalue observed in the spectrum approximated $\frac{13}{36}\pi^2$.

Convergence curves for 10 dominant eigenvalues of the mentioned hollow waveguide are plotted in figure 2.13. In the absence of an appropriate T/C method the so called low frequency instability would be observed in the *eigenvalue* 's and the convergence curves would not follow the typical algebraic (exponential) convergence characteristic presented in figure 2.13a (figure 2.13b). As apparent from figure 2.13, besides complete omission of zero valued spurious modes, the so called low frequency instability behavior of the FEM results has been completely suppressed. Without the implementation of a T/C method the accuracy of solutions would have been severely disturbed specially at the higher end of the plots in figure 2.13a. In both figure 2.13a and figure 2.13b, higher modes exhibit some kind of chaotic behavior at the lower end. The chaotic behavior disappears as the curves pass through the maximum mesh size (minimum interpolation order) dictated by

Nyquist 's [76] minimum sampling rate criterion. It must be emphasized that the maximum mesh size requirement is a function of frequency and material coefficient ϵ and μ and hence it is loser for the lower modes and tighter for higher modes.

While working with the matrix *eigenvalue* , it is observed that (when dealing with large problems) a σ slightly smaller than κ_{min}^2 (where κ_{min} represents the dominant eigenvalue) results in good convergence of the BKS solver. What happens with this choice of σ is that the shift and invert method starts by capturing the lowest (and thus most physically dominant) eigenvalues moving towards the higher ones.

2.9.2 Partially Loaded Waveguide

One of the advantages of Nedelec basis is that it allows for waveguide modal analysis without a need for a three dimensional field formulation. Using the basis, tangential field continuity is preserved while normal field components are allowed to be discontinuous at material interfaces. However, accurate and stable implementation of such modal analysis still depends on elimination of the unwanted gradient modes. The example discussed in this section involves a rectangular waveguide with the following cross sectional specifications:

- width $0 \leq x \leq a = 3$.
- height $0 \leq y \leq b = 1.5$.
- lower one third $0 \leq y \leq h = 0.5$ loaded with material $\epsilon_{r1} = 1, \mu_{r1} = 3$.
- upper two thirds $h = 0.5 \leq y \leq b = 1.5$ loaded with material $\epsilon_{r2} = 1, \mu_{r2} = 1$.
- PEC conditions assumed on the periphery.

For TE_{mn}^y modes, the radial cut-off frequency ω_c can be sought by solving the nonlinear (2.62) [75].

$$\begin{aligned}\epsilon_{r1}\sqrt{\omega_c^2\mu_{r2}\epsilon_{r2}-\left(\tfrac{m\pi}{a}\right)^2}\cot\left((b-h)\sqrt{\omega_c^2\mu_{r2}\epsilon_{r2}-(\tfrac{m\pi}{a})^2}\right)=\\ \epsilon_{r2}\sqrt{\omega_c^2\mu_{r1}\epsilon_{r1}-\left(\tfrac{m\pi}{a}\right)^2}\cot\left(h\sqrt{\omega_c^2\mu_{r1}\epsilon_{r1}-(\tfrac{m\pi}{a})^2}\right)\end{aligned} \tag{2.62}$$

Infinite number of solutions exist for each $m \geq 1$ in (2.62). The subscript n in TE^y_{mn} refers to the ascending order in which the mentioned roots are located on the real line. Table 2.5 lists the solutions for the first eight modes of (2.62) for the substitution $\kappa^2 = \epsilon_0\mu_0\omega_c^2$. These values can be directly compared to those found by FEM modal analysis. Figure 2.14 plots the h and p-convergence curves for the first 20 dominant modes in the partially loaded waveguide. Expectedly, no spurious modes were observed among the solutions due to the deployment of the $\mathcal{O}$ projection operator in the Krylov solver process. Similar analysis as presented in section 2.9.1 apply to the convergence curves presented in this section.

$\kappa(TE^y_{mn})$	m,n	$\kappa(TE^y_{mn})$	m,n
3.5943293313297	11	25.476300653105	13
6.0650093296508	21	28.055709664004	23
9.7105624257147	31	28.279756555185	71
12.421178219114	12	32.331619462127	52
13.886524562807	41	32.349023327324	33
14.451827595164	22	34.122452463520	81
18.186348755673	32	38.275190851237	43
18.287833863322	51	40.610458001668	91
23.032547204039	61	42.214180784259	62
24.120106962044	42	45.673879465019	53

Table 2.5: Square wavenumber κ^2 of the first 20 cut-off frequencies for TE^y_{mn} modes in the partially loaded waveguide obtained by numerical solution of (2.62).

2.9.3 Magnetostatic Boundary Value Problem

The *magnetostatic* example presented here considers the $2D$ magnetic field solution of an infinitely long hexagonal cross sectioned solenoid. The FEM process involves the solution of the weak form of the magnetic vector potential $\mathcal{A}$ where $\mathcal{B}$ (or the magnetic flux density) equals $\nabla \wedge \mathcal{A}$. Figure 2.16 schematically plots the problem geometry. The magnitude of the electric current density $|\mathcal{J}|$, as depicted in figure 2.16, is assumed to follow a second order function that vanishes on the inner and the outer boundaries of the current carrying region. Hence, $|\mathcal{J}|$ attains its

maximum between the two boundaries. The exterior boundary of the problem domain is assumed to terminate on a perfect magnetic conductor which imposes the $\hat{\mathbf{n}} \wedge \frac{1}{\mu}\nabla \wedge \mathcal{A} = 0$ condition. In this example, a uniform magnetic permeability μ_0 (free space) is assumed. The final matrix form of the problem, given in (2.63), is the same as (2.58).

$$\mathsf{A}\,\mathsf{x} = \mathsf{y}, \quad \mathsf{y} = \mathsf{B}\,\mathsf{J} \tag{2.63}$$

In (2.63), J is a vector which is obtained by interpolating the desired current density over the FE basis. It should be noted that the FEM matrices A and B are identical to those of (2.59) and (2.60), except that, μ_r is replaced by $\mu_0\mu_r$ in (2.59) and ϵ_r is dropped from (2.60).

Due to the discussed singularities of A, numerical solution of (2.63) is generally not possible unless the right hand side of the (2.63) is confined to the range-space of A. This can be achieved by applying the projection operator $\mathcal{O}$ to the right hand side of (2.63) and solving $\mathsf{A}\,\mathsf{x} = \mathcal{O}\,\mathsf{y}$ instead of $\mathsf{A}\,\mathsf{x} = \mathsf{y}$. In this example, fourth order-complete elements were used. This gave rise to 2190 Degree(s) of Freedom (DoF) for the problem (see figure 2.15). With the application of the mentioned projection operator, a *conjugated gradient squared (CGS)* solver took 260 iterations (approximately 0.2 seconds on a PC with a 3.00 GHz Pentium D) to reach the accuracy goal of 10^{-14}. Expected from the earlier analysis, attempting to solve the matrix problem without applying the projection operator resulted in numerical breakdown. Figure 2.18 plots the curl of the solved vector magnetic potential $\mathcal{A}$. The magnetic flux density increases over the current carrying region (inwards). In the current carrying region, since the current density follows a parabolic function (see figure 2.16), the flux density follows a third order curve along the radial direction. The magnetic flux density, however, is constant within the internal current free region. The calculated magnetic flux density thus complies with analytical predictions utilizing problem symmetry and *Ampere* 's law. Figure 2.17 depicts h and p convergence curves for the mentioned problem. The fact that the curves is not disturbed by the increase in the number of Degree(s) of

Freedom (DoF) indicates that the singularities of the LHS matrix of (2.59) are effectively handled by the projection operator $\mathcal{O}$.

2.10 Conclusion

The presented approach can accurately model (and handle) the singularities of FE matrices in EM FE analysis with HO spectral elements of the Nedelec type [12, 65, 66]. Utilization of the dual grid results in a local-to-global matrix assemblage method for the constraints matrix. The required constraints matrix is directly assembled from a constant element matrix bypassing the need for evaluation of discrete integration and gradient operators [58]. The proposed element matrix is proved to be independent from element geometry and metric properties (including curved elements). Thus, it is concluded that the proposed HO constraints matrix is solely determined by the topology of the FE mesh.

It is worth mentioning that a comparison against the method of direct imposition of the divergence condition requires developing FEM codes that involve in a mixed (H^1 and $H(\nabla\wedge)$) FE formulations and repeated solutions of auxiliary Poisson's problems [67, 68]. This has not been addressed in this article. It is generally preferable if such formulations can be avoided [58]. The approach presented here should rather be considered as an extension to the [58] while the following features make it distinguished from [58]'s approach.

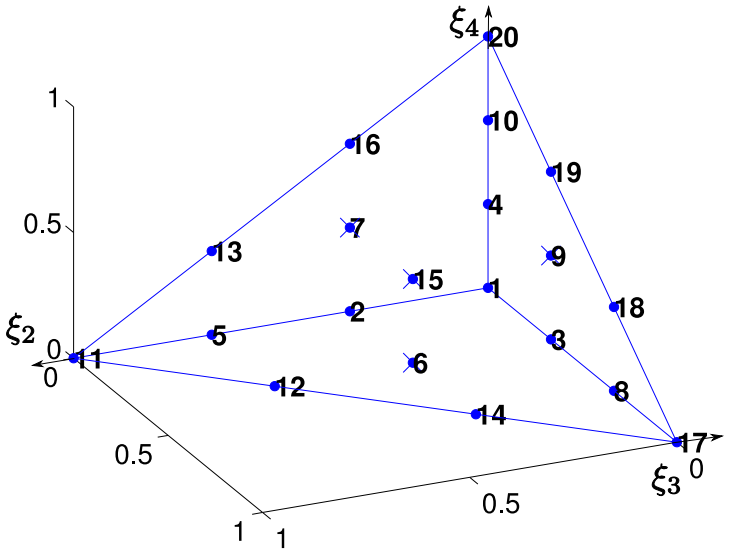
1. It works for HO spectral basis.
2. The analysis is based on the functional properties of the involved FE spaces rather than the physical imposition of the divergence condition.
3. It is more efficient while assembling the constraints matrix.
4. It is based on the dual grid.

Despite the interpolatory orientation of the work, the method can be applied to cases with

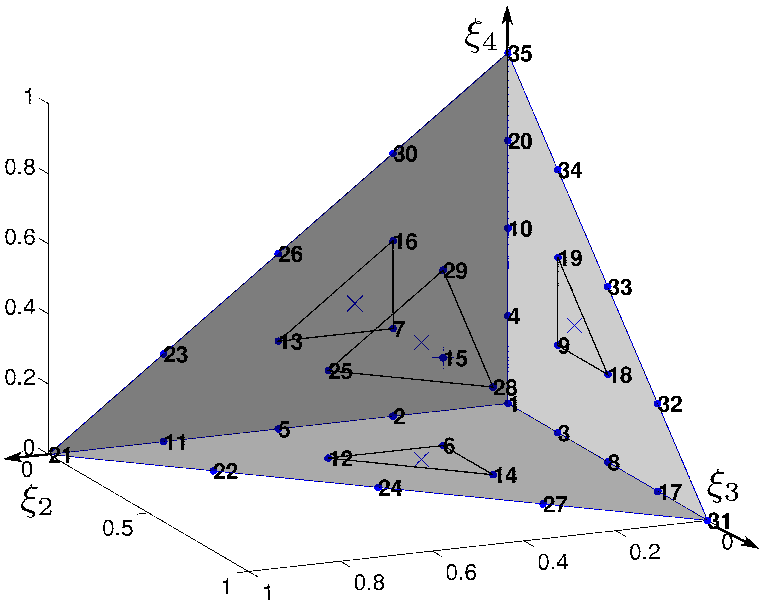
hierarchical elements since only a first order part of the hierarchical basis requires construction of such constraint equations [56]. As a natural property of an spectral method, the presented approach is unlikely to be extended to p-nonuniform cases. Yet, the hierarchical version of the dual-grid based approach would have no limitations in dealing with p-nonuniformity.

The effect of the choice of the dual tree on the sparsity of the constraints matrix is explicitly formulated. This allows for efficient preallocation of the sparse structure of (the constraints matrix) before its actual assemblage. It is possible to utilize this formulation for optimal choice of the dual tree that should lead to substantial savings in computer memory and CPU flops consumption. Included numerical examples verify the proposed methodology. Furthermore, it is observed that the size of the constraints matrix is affected by the choice of the dual tree. In this regard, future related work must be focused on obtaining optimal spanning trees.

It is theoretically shown that the proposed dual-grid based approach can be extended to $3D$ cases with simplex shaped elements. Extension of the dual-gird based method to hexahedra, quadrangles and prisms (and mixed meshes) depends on the possibility of constructing the basis with rotational symmetries analogous to those developed in this article. Without such developments, the dual-grid based approach cannot be used for cases with hexahedral, quadrangular and prismic meshes for which the conventional approach (which directly enforces the divergence condition) has been successfully implemented.

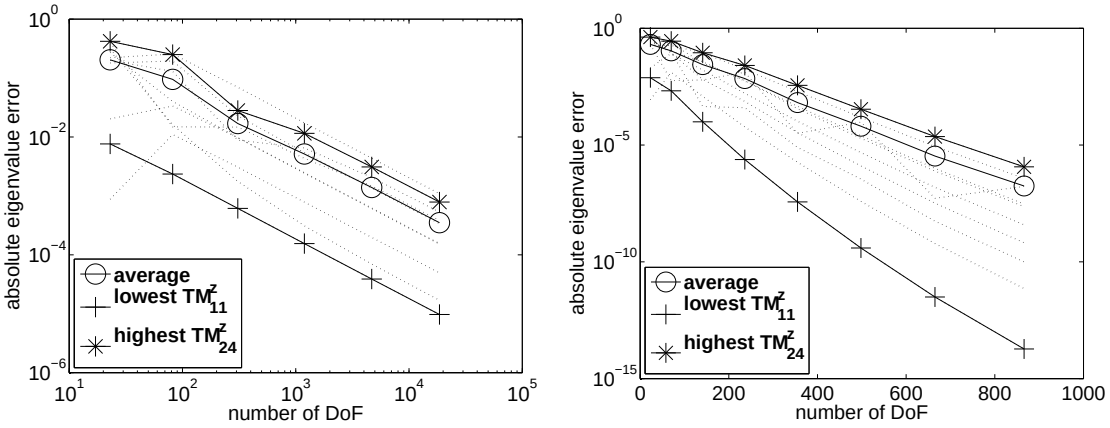


(a) H^1 nodal set for polynomial order $p = 3$.



(b) $H(\nabla\wedge)$ nodal set for polynomial order $p = 2$. Note that no DoF is associated to the vertex nodes. there are 3 nodes per each face and one node inside. For better clarity, the face nodes are tied together with the aid of triangles.

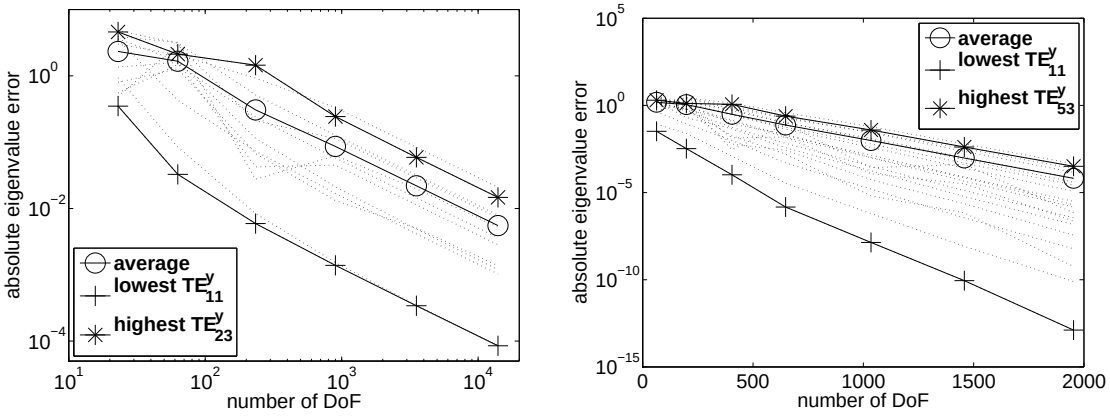
Figure 2.12: Configuration of interpolation nodal sets for 3D tetrahedral elements.



(a) h-convergence diagram showing 10 dominant eigenvalues' absolute error. The corresponding h values are $\{1, 0.5, 0.25, 0.125, 0.0625, 0.03125\}$.

(b) p-convergence diagram showing 10 dominant eigenvalues' absolute error. The corresponding p values are $\{0, 1, 2, 3, 4, 5, 7\}$.

Figure 2.13: h and p convergence diagrams showing absolute error in calculated eigenvalues for the hollow waveguide. The modes are labeled in ascending order.



(a) h-convergence for 20 dominant eigenvalue's, $h \in \{1, 0.5, 0.25, 0.125, 0.0625, 0.03125\}$.

(b) p-convergence for 20 dominant eigenvalue's, $p \in \{0, 1, 2, 3, 4, 5, 6\}$.

Figure 2.14: Convergence for the partially loaded waveguide. Modes in ascending order.

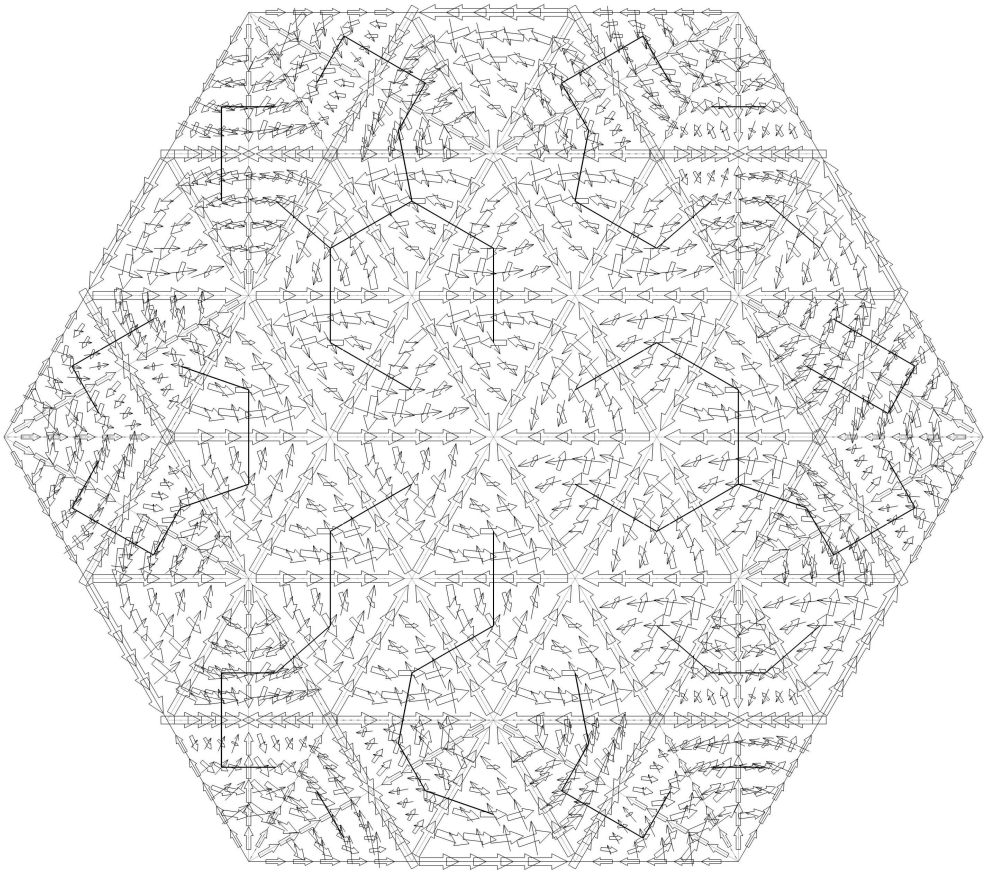


Figure 2.15: Problem mesh and associated 4^{th}-order DoF for the hexagonal solenoid problem. The lines connecting the element centers are the dual tree lines except that lines connecting the tree elements to the exteriority are not plotted

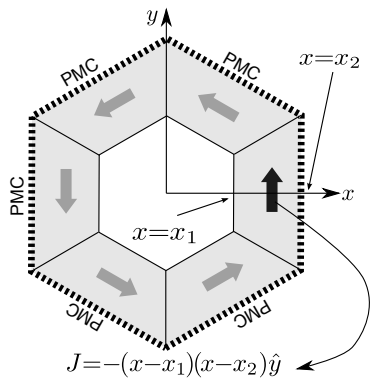
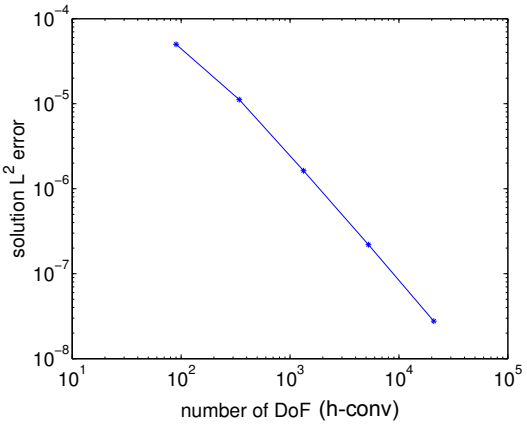
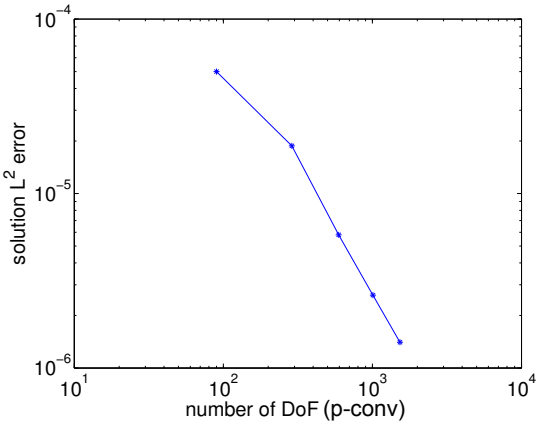


Figure 2.16: Problem geometry for the hexagonal solenoid problem. The shaded region indicates presence of electrical current density $\jmath$. For the region with the darker arrow we have $\jmath = -(x-x_1)(x-x_2)\hat{y}$.



(a) h-convergence diagram showing magnetic field solution's L^2 error versus number of DoF. The corresponding h values are $\{1, 0.5, 0.25, 0.125, 0.0625\}$.



(b) p-convergence diagram showing magnetic field solution's L^2 error versus number of DoF. The corresponding p values are $\{0, 1, 2, 3, 4\}$.

Figure 2.17: h and p convergence diagrams for magnetic field solution in the hexagonal solenoid.

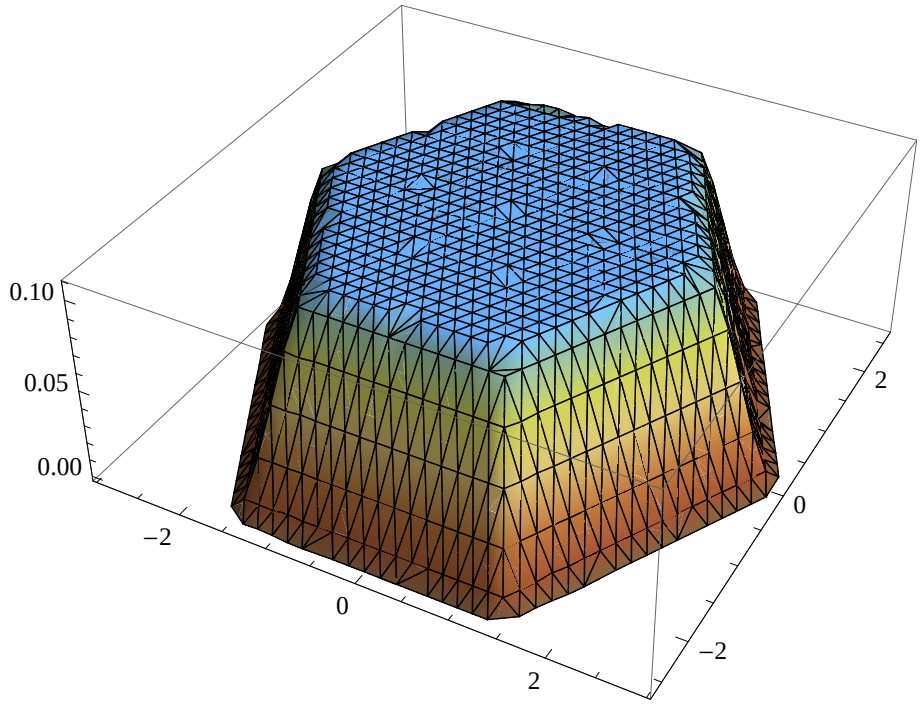


Figure 2.18: Magnetic flux $\mathcal{B}$ for the solenoid problem. Irregularities are the side effect the first order visualization of the 4^{th} order-complete solution.

Chapter 3

Finite/Infinite Elements in H^1

3.1 Introduction

Evaluation of the weak/variational form is a essential part of the FEM procedure. The weak form is a bilinear form defined as an integral and over the entire problem domain and its bundaries. On the other hand, almost all real world problems are acctually defined over an infinite/untruncated domain. Yet, truncation of the actual problem domain with artificial boundary conditions is common practice specially when the nature of the problem is local. For example, in cases such as Poisson's equation, the actual problem domain is often truncated with homogeneous Neumann boundary conditions imposed at a boundary which is sufficiently apart from the actual region of interest. Such truncation of problem domain is coherent with our physical intuition: sufficiently apart objects have negligible effect on our actual physical experiment. Such truncation techniques, however, are not applicable to wave phenomena as wave energy is reflected and redirected into the domain of interest resulting in a physically inconsistent model. For example, a radiation source cannot be enclosed in a PEC sphere of any radius to reproduce the free space radiation effect and that's why anechoic chambers were ever invented. Consequently, for problems involving wave phenomena, special truncation techniques have been devised. The current chapter is dedicated to a discussion on the IE method is one of the abovementioned techniques.

The governing idea behind IEs is to discretize a part of the problem domain that approaches infinity using elements which's dimension extend to infinity along one direction. Moreover, the free space solution of the wave problem is incorporated to the functional expressions reconstructing the outgoing wave behavior(functions) of the wave phenomenon. Similar to FEs, IE matrices can be assembled from local matrices solely determined by individual IEs. It is worth mentioning that IE were first devised for static problems and their application is by no means limited to wave phenomena.

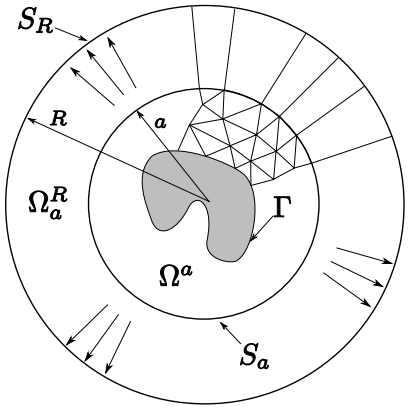


Figure 3.1: Truncation of FE mesh with infinite elements. The shaded region located in the center is considered to be a rigid body.

3.2 Problem Definition

A consistent formulation for a general *Helmholtz* radiation and problem will be developed in this section. As depicted in figure 3.1 the problem domain is partitioned into two parts:

1. Ω^a which includes all possible geometrical and material complexities except for the interior of rigid bodies. It is further required that f has a compact support inside Ω^a. In other words, f must be entirely zero outside Ω^a.

2. Ω_a^R which must only include free space and will be extended to infinity as $R \to \infty$.

Furthermore, it is assumed that the solution is the superposition of an incident u^{inc} and a reaction/scattered u wave. Hence, the PDE formulation of the problem is stated as (3.1) where f

represents any existing independent sources and d stands for the number of spatial dimensions.

$$\begin{cases} (-\nabla \cdot \epsilon \nabla - k^2)(u + u^{inc}) = f \ in \ \mathbb{R}^d, d = 2, 3 \\ \\ \text{boundary conditions satisfied for } (u + u^{inc}) \end{cases} \tag{3.1}$$

The incident wave function u^{inc} is assumed to be a free space wave function such as a plane wave. Thus, u^{inc} automatically satisfies $(-\nabla \cdot \epsilon \nabla - k^2)u^{inc} = 0$. Hence, it remains to us to find u such that (3.2) is fulfilled.

$$\begin{cases} (-\nabla \cdot \epsilon \nabla - k^2)u = f \ in \ \mathbb{R}^d, d = 2, 3 \\ \\ \text{boundary conditions satisfied for } (u + u^{inc}) \end{cases} \tag{3.2}$$

3.3 *Sommerfeld* Radiation Condition

Setting aside the incident wave u^{inc}, it is clear that u is the part of the field that originates inside Ω^a and radiates into infinite space. Thus, the physical solutions of Helmholtz's scalar wave equation (3.2) must satisfy the radiation condition (3.3) in which an $e^{\jmath\omega t}$ time dependence is assumed.

$$\alpha \triangleq \frac{d-1}{2}, \ \lim_{|x| \to \infty} |x|^{\alpha} (\frac{\partial}{\partial |x|} + \jmath k) u = 0 \tag{3.3}$$

In (3.3), α is a problem dependent parameter that equals to $1/2$ and 1 respectively for two and three dimensional problems. The radiation condition has been thoroughly discussed in scattering and radiation texts. Thus, I suffice to quote *Sommerfeld* 's original interpretation:'the sources must be sources, not sinks of energy. The energy which is radiated from the sources must scatter to infinity; no energy may be radiated from infinity into ... the field.'[77].

Now, suppose u is the solution of a scalar Helmholtz scattering problem in Ω_a^R. Sommerfeld radiation condition states that $\frac{\partial u}{\partial r} + \jmath k u$ must be L^2 integrable in Ω_a^R:

$$\forall R \in \mathbb{R} \bigcup \{\infty\}, (\frac{\partial u}{\partial |x|} + \jmath k u) \in L^2(\Omega_a^R), \quad \Omega_a^R \triangleq \{x \in \mathbb{R}^n | a < |x| < R\} \tag{3.4}$$

At the same time, for any function $w \in L^2(\Omega_a^R)$, one observes that:

$$\lim_{R\to\infty} \int_{r=a}^{R} \int_{S_r} |w|^2 dS_r dr < \infty \Rightarrow \lim_{r\to\infty} \int_{S_r} |w|^2 dS_r = 0 \tag{3.5}$$

Hence, from (3.4) and (3.5), (3.6) is concluded.

$$\lim_{r\to\infty} \int_{S_r} |\frac{\partial u}{\partial r} + \jmath k u|^2 dS_r = 0 \tag{3.6}$$

This is the form of the Sommerfeld radiation condition that will be used throughout this work. The F/IE approach partitions the problem domain into two regions: Ω^a and Ω_a^R. Throughout this chapter, Ω^a and Ω_a^R are respectively regarded as the FE and the IE domain.

3.4 Weak Formulation

Following the partition of the problem domain into Ω^a and Ω_a^R regions, the weak form of the problem is split into I_{fin} and I_{inf} respectively. Both I_{fin} and I_{inf} are obtained from (3.2). Indeed, the complete weak form is obtained by the superposition of the I_{fin} and I_{inf} in which the fin and inf subscripts indicate the respective use of 'finite' and 'infinite' elements for the final discretization. It is thus clear that I_{fin} and I_{inf} are respectively associated to Ω^a and Ω_a^R. First, the evaluation of I_{fin} is discussed.

3.4.1 Evaluation of I_{fin}

Going back to the problem model in figure 3.1 it is observed that the solution has to satisfy $\nabla(u^{inc} + u).\hat{n}|_\Gamma = 0$ as Γ represents a rigid body surface. However, to keep the formulation general we specify boundary condition (3.7) on Γ and emphasize that for a rigid scatterer surface Γ, the system is excited by $g = -\epsilon \nabla u^{inc} \cdot \hat{n}$ and $\gamma = 0$ on Γ.

$$\epsilon \nabla u^{total} \cdot \hat{n} + \gamma u^{total} = g \tag{3.7}$$

Hence, multiplying (3.2) by $\bar{v}$ and integrating the whole expression over Ω^a and incorporating (3.7) yields (3.8). Derivation of (3.8) involves substitution of the vector identity (3.9) and the Green's identity (3.10).

$$I_{fin}: \int_{\Omega^a} (\nabla\bar{v}\cdot\epsilon\nabla u - \bar{v}k^2 u)\, d\Omega^a + \int_{\Gamma} (\bar{v}\gamma u)\, d\partial\Omega^a = \int_{\Omega^a} \bar{v} f\, d\Omega^a + \int_{S_a} \bar{v}\epsilon\nabla u\cdot\hat{n}\, d\partial\Omega^a + \int_{\Gamma} \bar{v} g\, d\partial\Omega^a \tag{3.8}$$

$$\nabla\cdot(v\epsilon\nabla u) = \nabla v\cdot\epsilon\nabla u + v\nabla\cdot\epsilon\nabla u \tag{3.9}$$

$$\int_{\Omega} \nabla\cdot w\, d\Omega = \int_{\partial\Omega} w\cdot\hat{n}\, d\partial\Omega \tag{3.10}$$

3.4.2 Evaluation I_{inf}

Our intention is to implement the F/IE method to a 2D scalar Helmholtz problem with a circular truncation of the FE mesh at an enclosing circular boundary S_a beyond which the IEs extend into infinity. As depicted in figure 3.1, the IEs radially extend from S_a to a second circle S_R for which the a final extension to infinity is sought by means of imposing $R \to \infty$. The very first step is the evaluation of the weak form of the problem for the exterior domain Ω_a^R. Similar to the previous section, the weak form appears as (3.11) for which the $R \to \infty$ limit will be imposed later. For future convenience, an extra $\int_{S_R} \bar{v}\jmath k u\, d\partial\Omega^a$ has been added to both sides of the equation. Hence, we have:

$$I_{inf}: \int_{\Omega_a^R} (\nabla\bar{v}\cdot\epsilon\nabla u - \bar{v}k^2 u)\, d\Omega^a + \int_{S_R} \bar{v}\jmath k u\, d\partial\Omega^a = \int_{S_a} \bar{v}\epsilon\nabla u\cdot\hat{n}\, d\partial\Omega^a + \int_{S_R} \bar{v}(\epsilon\nabla u\cdot\hat{n} + \jmath k u)\, d\partial\Omega^a \tag{3.11}$$

However, the RHS terms of (3.11) will, later on, be ignored because:

1. The first term in the RHS cancels off with the same term in I_{fin} evaluated for the interior domain Ω_a^R.

2. Due to the Sommerfeld radiation condition (3.6) and the *Cauchy -Schwarz* inequality in *L^2 functional space (L^2)* , the second term on the RHS of (3.11) must vanishes when $R \to \infty$. Note that in (3.11), $\nabla u \cdot \hat{n}|_{S_R} = \frac{\partial u}{\partial \xi_1}$ where ξ_1 is the radius from the origin of the coordinate system. Also, since Ω_a^R models the infinite free space into which the wave energy is radiated, it is obvious that ϵ equals to the identity tensor $\mathsf{I}_{d\times d} : \mathbb{R}^2 \to \mathbb{R}^2$ in Ω_a^R. Thus we take the limit of I_{inf} as:

$$\lim_{R\to\infty} I_{inf} : \int_{\Omega_a^R} (\nabla \bar{v} \cdot \epsilon \nabla u - \bar{v} k^2 u)\, d\Omega^a + \int_{S_R} \bar{v} \jmath k u \, d\partial\Omega^a = \int_{S_a} \bar{v} \epsilon \nabla u \cdot \hat{n} \, d\partial\Omega^a \tag{3.12}$$

Due to the circular shape of the FE truncation boundary S_R, it is natural to choose a transformation into polar coordinates (ξ_1, ξ_2), i.e. (3.13).

$$\begin{cases} x_1 = \xi_1 \; a \; cos(b\xi_2 + c) \\ \\ x_2 = \xi_1 \; a \; sin(b\xi_2 + c) \end{cases} \tag{3.13}$$

The forward and backward Jacobian matrices associated to this transformation are provided in (3.14) and (3.15).

$$\mathsf{J} = \left[\frac{\partial x_i}{\partial \xi_j}\right] = \begin{bmatrix} a\, cos(c + b\xi_2) & -a\; b\; \xi_1 sin(c + b\xi_2) \\ \\ a\, sin(c + b\xi_2) & a\; b\; \xi_1 cos(c + b\xi_2) \end{bmatrix} \tag{3.14}$$

$$\mathsf{J}^{-1} \triangleq \mathsf{K} = \left[\frac{\partial \xi_i}{\partial x_j}\right] = \frac{1}{\det(\mathsf{J})} \begin{bmatrix} a\; b\; \xi_1 cos(c + b\xi_2) & a\; b\; \xi_1 sin(c + b\xi_2) \\ \\ -a\; sin(c + b\xi_2) & a\; cos(c + b\xi_2) \end{bmatrix} \tag{3.15}$$

Now, since $\nabla_x \bullet \triangleq [\frac{\partial \bullet}{\partial x_i}] = \mathsf{K}^T \times \nabla_\xi \bullet$ and since $\epsilon = \mathsf{I}_{d\times d}$, one deduces that:

$$\nabla_x u \cdot \epsilon \nabla_x \bar{v} = \begin{bmatrix} \partial_{\xi_1} u & \partial_{\xi_2} u \end{bmatrix} \times \mathsf{K} \times \mathsf{K}^T \times \begin{bmatrix} \partial_{\xi_1} \bar{v} \\ \partial_{\xi_2} \bar{v} \end{bmatrix} = \begin{bmatrix} \partial_{\xi_1} u & \partial_{\xi_2} u \end{bmatrix} \times \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{a^2 b^2 \xi_1^2} \end{bmatrix} \times \begin{bmatrix} \partial_{\xi_1} \bar{v} \\ \partial_{\xi_2} \bar{v} \end{bmatrix} \tag{3.16}$$

As we intend to evaluate I_{inf} of (3.12), (3.15) and (3.16) are used to yield the following expression for I_{inf}:

$$I_{inf}: \int_{\xi_2=0}^{1} \int_{\xi_1=1}^{\frac{R}{a}} \left(\frac{\partial \bar{v}}{\partial \xi_1} \frac{\partial u}{\partial \xi_1} + \frac{1}{b^2 \xi_1^2} \frac{\partial \bar{v}}{\partial \xi_2} \frac{\partial u}{\partial \xi_2} - k^2 a^2 \bar{v} u \right) b \xi_1 d\xi_1 d\xi_2 + \jmath k \int_{\xi_2=0}^{1} \bar{v} u b a \xi_1 |_{\xi_1=\frac{R}{a}} d\xi_2 = 0 \tag{3.17}$$

Note that for (3.17) to cover the entire Ω_a^R domain, the b coefficient in (3.13) must be equal to 2π. However, arbitrary values of b and c in (3.13) allow for evaluation of I_{inf} over individual IEs the combination of which covers Ω_a^R.

3.4.2.1 First Variable Substitution

From this point on, a sequence of variable transformations are introduced. The rationale behind the choice of these transformations will be discussed after the final formulation for I_{inf} is obtained. The very first variable transformation introduced would be (3.18).

$$\begin{cases} u = e^{-\jmath k a (\xi_1 - 1)} U \\ v = e^{-\jmath k a (\xi_1 - 1)} V \end{cases} \tag{3.18}$$

Equation (3.18) results in the transformations listed in (3.19) through (3.23).

$$\frac{\partial u}{\partial \xi_1} = e^{-\jmath ka(\xi_1 - 1)} \left(\frac{\partial U}{\partial \xi_1} - \jmath kaU \right) \tag{3.19}$$

$$\frac{\partial \bar{v}}{\partial \xi_1} = e^{+\jmath ka(\xi_1 - 1)} \left(\frac{\partial \bar{V}}{\partial \xi_1} + \jmath ka\bar{V} \right) \tag{3.20}$$

$$\frac{\partial u}{\partial \xi_1}\frac{\partial \bar{v}}{\partial \xi_1} = \left(\frac{\partial U}{\partial \xi_1} - \jmath kaU \right) \left(\frac{\partial \bar{V}}{\partial \xi_1} + \jmath ka\bar{V} \right) \tag{3.21}$$

$$\frac{\partial u}{\partial \xi_2}\frac{\partial \bar{v}}{\partial \xi_2} = \frac{\partial U}{\partial \xi_2}\frac{\partial \bar{V}}{\partial \xi_2} \tag{3.22}$$

$$u\bar{v} = U\bar{V} \tag{3.23}$$

Thus, I_{inf} transforms to (3.24).

$$\begin{aligned} I_{inf} : \int_{\xi_2=0}^{1} \int_{\xi_1=1}^{\frac{R}{a}} \left(\frac{\partial \bar{V}}{\partial \xi_1}\frac{\partial \bar{U}}{\partial \xi_1} + \jmath ka \left(\frac{\partial U}{\partial \xi_1}\bar{V} - \frac{\partial \bar{V}}{\partial \xi_1}U \right) + \frac{1}{b^2 {\xi_1}^2}\frac{\partial \bar{V}}{\partial \xi_2}\frac{\partial U}{\partial \xi_2} \right) b\xi_1 d\xi_1 d\xi_2 \\ + \jmath k \int_{\xi_2=0}^{1} \bar{V}Uba\xi_1|_{\xi_1=\frac{R}{a}} d\xi_2 = 0 \end{aligned} \tag{3.24}$$

3.4.2.2 Second Variable Substitution

Equation (3.25) presents the second variable transformation that would be imposed. In (3.25), α is a real constant that respectively assumes the values of $\frac{1}{2}$ or 1 depending on whether a 2D or 3D problem is involved.

$$\begin{cases} U = {\xi_1}^{-\alpha} S \\ V = {\xi_1}^{-\alpha} T \end{cases} \tag{3.25}$$

Imposing (3.25), the transformations listed in (3.26) through (3.31) occur. Incorporating these transformations into (3.24) one ends up with (3.32).

$$\frac{\partial U}{\partial \xi_1} = -\alpha {\xi_1}^{-\alpha-1} S + {\xi_1}^{-\alpha} \frac{\partial S}{\partial \xi_1} = {\xi_1}^{-\alpha} \left(\frac{\partial S}{\partial \xi_1} - \frac{\alpha}{\xi_1} S \right) \tag{3.26}$$

$$\frac{\partial \bar{V}}{\partial \xi_1} = -\alpha {\xi_1}^{-\alpha-1} \bar{T} + {\xi_1}^{-\alpha} \frac{\partial \bar{T}}{\partial \xi_1} = {\xi_1}^{-\alpha} \left(\frac{\partial \bar{T}}{\partial \xi_1} - \frac{\alpha}{\xi_1} \bar{T} \right) \tag{3.27}$$

$$\frac{\partial U}{\partial \xi_2} = {\xi_1}^{-\alpha} \frac{\partial S}{\partial \xi_2} \tag{3.28}$$

$$\frac{\partial \bar{V}}{\partial \xi_2} = {\xi_1}^{-\alpha} \frac{\partial \bar{T}}{\partial \xi_2} \tag{3.29}$$

$$\left(\frac{\partial U}{\partial \xi_1} \bar{V} - \frac{\partial \bar{V}}{\partial \xi_1} U \right) = {\xi_1}^{-2\alpha} \left(\frac{\partial S}{\partial \xi_1} \bar{T} - \frac{\partial \bar{T}}{\partial \xi_1} S \right) \tag{3.30}$$

$$\frac{\partial U}{\partial \xi_1} \frac{\partial \bar{V}}{\partial \xi_1} = {\xi_1}^{-2\alpha} \left(\frac{\partial S}{\partial \xi_1} - \frac{\alpha}{\xi_1} S \right) \left(\frac{\partial \bar{T}}{\partial \xi_1} - \frac{\alpha}{\xi_1} \bar{T} \right) \tag{3.31}$$

$$\begin{aligned} I_{inf} : & \int_{\xi_2=0}^{1} \int_{\xi_1=1}^{\frac{R}{a}} \left(\left(\frac{\partial \bar{T}}{\partial \xi_1} - \frac{\alpha \bar{T}}{\xi_1} \right) \left(\frac{\partial S}{\partial \xi_1} - \frac{\alpha S}{\xi_1} \right) + \jmath ka \left(\frac{\partial S}{\partial \xi_1} \bar{T} - \frac{\partial \bar{T}}{\partial \xi_1} S \right) + \frac{1}{b^2 \xi_1^2} \frac{\partial \bar{T}}{\partial \xi_2} \frac{\partial S}{\partial \xi_2} \right) b d\xi_1 d\xi_2 \\ & + \jmath kab \int_{\xi_2=0}^{1} \bar{T} S|_{\xi_1 = \frac{R}{a}} d\xi_2 = 0 \end{aligned} \tag{3.32}$$

Applying by-parts integration onto the second RHS term in (3.32) yields the following expression for I_{inf}:

$$\begin{aligned} & \int_{\xi_2=0}^{1} \left(\int_{\xi_1=1}^{\frac{R}{a}} \left(\left(\frac{\partial \bar{T}}{\partial \xi_1} - \frac{\alpha \bar{T}}{\xi_1} \right) \left(\frac{\partial S}{\partial \xi_1} - \frac{\alpha S}{\xi_1} \right) + 2\jmath ka \frac{\partial S}{\partial \xi_1} \bar{T} + \frac{1}{b^2 {\xi_1}^2} \frac{\partial \bar{T}}{\partial \xi_2} \frac{\partial S}{\partial \xi_2} \right) b \, d\xi_1 - \jmath kab \bar{T} S|_{\xi_1=1}^{\frac{R}{a}} \right) d\xi_2 \\ & + \jmath kab \int_{\xi_2=0}^{1} \bar{T} S|_{\xi_1 = \frac{R}{a}} d\xi_2 = 0 \end{aligned} \tag{3.33}$$

If (3.33) is expanded and a second by-parts integration is imposed on the $-\frac{\alpha}{\xi_1} S \frac{\partial \bar{T}}{\partial \xi_1}$ term, (3.34) results. It must be emphasized here that we are intended to calculate $\lim_{R \to \infty} I_{inf}$ and not I_{inf} as it appears in (3.33). Therefore, at this stage, the $R \to \infty$ limit is taken which causes the last integral term on the RHS of (3.34) to vanish.

$$I_{inf}: \int_{\xi_2=0}^{1}\left(\int_{\xi_1=1}^{\frac{R}{a}}\left(\frac{\partial\bar{T}}{\partial\xi_1}\frac{\partial S}{\partial\xi_1}+\frac{\alpha(\alpha-1)}{{\xi_1}^2}\bar{T}S+2\jmath ka\frac{\partial S}{\partial\xi_1}\bar{T}+\frac{1}{b^2{\xi_1}^2}\frac{\partial\bar{T}}{\partial\xi_2}\frac{\partial S}{\partial\xi_2}\right)bd\xi_1\right)d\xi_2$$
$$+(\jmath ka+\alpha)b\int_{\xi_2=0}^{1}\bar{T}S|_{\xi_1=1}d\xi_2-\frac{\alpha ab}{R}\int_{\xi_2=0}^{1}\bar{T}S|_{\xi_1=\frac{R}{a}}d\xi_2=0 \tag{3.34}$$

3.4.2.3 The last change of variable

The last change of variable imposed here would be $\zeta=\frac{1}{\xi_1}$. Below are the transformations that follow as the consequence.

$$d\xi_1=-\frac{1}{\zeta^2}d\zeta \tag{3.35}$$
$$\frac{\partial S}{\partial\xi_1}=-\zeta^2\frac{\partial S}{\partial\zeta} \tag{3.36}$$
$$\frac{\partial\bar{T}}{\partial\xi_1}=-\zeta^2\frac{\partial\bar{T}}{\partial\zeta} \tag{3.37}$$

Incorporating (3.35) through (3.37) in (3.34) leads to (3.38) in which S and T are related to u and v by (3.39).

$$\lim_{R\to\infty} I_{inf}: \int_{\xi_2=0}^{1}b\left(\int_{\zeta=0}^{1}\left(\zeta^2\frac{\partial\bar{T}}{\partial\zeta}\frac{\partial S}{\partial\zeta}+\alpha(\alpha-1)\bar{T}S-2\jmath ka\frac{\partial S}{\partial\zeta}\bar{T}+\frac{1}{b^2}\frac{\partial\bar{T}}{\partial\xi_2}\frac{\partial S}{\partial\xi_2}\right)d\zeta\right)d\xi_2$$
$$+(\jmath ka+\alpha)b\int_{\xi_2=0}^{1}\bar{T}S|_{\zeta=1}d\xi_2 \tag{3.38}$$

$$\begin{cases} u=e^{\jmath ka(1-1/\zeta)}\zeta^\alpha S \\ v=e^{\jmath ka(1-1/\zeta)}\zeta^\alpha T \end{cases} \tag{3.39}$$

3.4.2.4 Discussion

The sequence of variable substitutions introduced in the previous sections does indeed have a physical meaning. In order to appreciate this, let's focus on the fact that outgoing time-harmonic cylindrical waves can be expanded using *Hankel* functions [75]:

$$\sum_{m\in\mathbb{Z}^+}(c_m cos(m\phi)+s_m sin(m\phi))H_m^{(2)}(k\xi_1) \tag{3.40}$$

Now, let us consider the asymptotic expansion of $H_m^{(2)}(k\xi_1)$ around ∞:

$$H_m^{(2)}(k\xi_1)=\frac{e^{-ik\xi_1}}{\sqrt{k\xi_1}}\sum_{i\in\mathbb{Z}^+}\frac{h_{mi}}{k^i\xi_1^i} \tag{3.41}$$

Combining (3.40) and (3.41), a general asymptotic form for outgoing time-harmonic cylindrical waves is obtained as:

$$\frac{e^{-ik\xi_1}}{\sqrt{k\xi_1}}\sum_{m,i\in\mathbb{Z}^+}(c_m cos(m\phi)+s_m sin(m\phi))\frac{h_{mi}}{k^i}\frac{1}{\xi_1^i} \tag{3.42}$$

With the above expansion of outgoing cylindrical waves taken into consideration, a second look at (3.38) and (3.39), leads to insightful understanding of (3.38): approximation of S and T functions of (3.38) using the usual H^1 polynomials in the (ζ,ξ_2) domain leads to expansions in ζ of the form (3.43). Take note that $\alpha=1/2$ in $2D$. The similarity in the ξ_1 dependence of (3.42) and (3.43) is by no means coincidental and indicates how the general forms of u and v expansions (as in (3.39)) allow for appropriate reconstruction of the required time-harmonic outgoing waves. This discussion can be analogously extended to $3D$ with $\alpha=1$ and spherical *Bessel* functions in the formulation of time-harmonic outgoing spherical waves.

$$e^{\jmath ka(1-1/\zeta)}\zeta^\alpha\sum_{i\in\mathbb{Z}^+}c_i\zeta^i=e^{\jmath ka}\frac{e^{-\jmath ka\xi_1}}{\xi_1^\alpha}\sum_{i\in\mathbb{Z}^+}c_i\frac{1}{\xi_1^i} \tag{3.43}$$

3.5 Discretization of I_{fin}

This section focuses on discretization of I_{fin} as formulated in (3.8). Technically speaking, the FE domain Ω^a is triangulated using rectilinear and curved elements. The rightmost RHS term in (3.8) will be dropped as it will cancel off with a similar term in I_{inf}. I_{fin} will be evaluated on individual elements and thus the individual element contributions to I_{fin} will be denoted by $I_{fin,e}$. The u, v, f and u^{inc} terms in (3.8) are approximated using H^1 polynomials and the integration is pulled-back onto a reference element $\triangle = (\xi_1, \xi_2)|0 \leq \xi_1 + \xi_2 \leq 1, \xi_1, \xi_2 \geq 0$. Due to the curvature of S_a as depicted in figure 3.1 there is a need for curved elements in the FE domain where FEs meets the IE domain. We shall return to evaluation of FE matrices over curved elements later on in section 3.5.1. Regardless of the rectilinear curvilinear nature of FEs, let $\{\alpha_i\}$ denote an appropriate H^1 Lagrangian basis defined over the reference element $\triangle$.

$$\begin{cases} v = \sum_i v_i \alpha_i(\xi_1, \xi_2) \\ u = \sum_j u_j \alpha_j(\xi_1, \xi_2) \end{cases} \tag{3.44}$$

Hence, (3.8) will be discretize to (3.45), from which the formulation of FE matrices is clear.

$$I_{fin,e} : \sum_{i,j} \bar{v}_i u_j \left(\int_{\Omega^a} (\nabla \bar{\alpha}_i \cdot \nabla \alpha_j - k^2 \bar{\alpha}_i \alpha_j) \, d\Omega^a \right) = \sum_{i,j} \bar{v}_i u_j \left(\int_{\Omega^a} \bar{\alpha}_i f_j \, d\Omega^a - \int_{\Gamma} \bar{\alpha}_i \nabla u_j^{inc} \cdot \hat{n} \, d\partial\Omega^a \right) \tag{3.45}$$

3.5.1 Efficient Evaluation of FE Matrices

In this section we shall discuss a technique is introduced for the evaluation of FE matrices. The method presented here can be equally applied to evaluation of curvilinear or rectilinear FE matrices. In chapter 5, this method is further extended to a more general case with $H(\nabla\wedge)$ elements.

What will be developed in this section is based on the following assumptions:

1. Element barycentric coordinates (reference element) are denoted by $\{\xi_i\}$ and actual physical coordinates is denoted by $\{x_i\}$.

2. Further assume that a Jacobian matrix J is associated to the transformation between the two coordinate systems.

$$\mathsf{J} = \left[\frac{\partial x_i}{\partial \xi_j}\right] \tag{3.46}$$

$$\mathsf{K} = \left[\frac{\partial \xi_i}{\partial x_j}\right] \tag{3.47}$$

   Indeed $\mathsf{J} = \mathsf{K}^{-1}$. Furthermore, we will denote the $H^1(\Omega)$ conforming element basis functions with $\{\alpha_i\}$ and the $H^1(\Omega)$ conforming element Lagrangian basis functions by $\{\ell_i\}$.

3. If we define ∇_ξ and ∇_x as $\nabla_\xi = [\partial \bullet / \partial \xi_i]^\mathsf{T}$ and $\nabla_x = [\partial \bullet / \partial x_i]^\mathsf{T}$, then the following transformation rules apply to any sufficiently smooth function u in ξ (or x).

$$\nabla_\xi u = \mathsf{J}^\mathsf{T} \times \nabla_x u \tag{3.48}$$

$$\nabla_x u = \mathsf{K}^\mathsf{T} \times \nabla_\xi u \tag{3.49}$$

4. Last we need to clarify that the actual integration for FE matrices is defined on the physical element domain denoted by $\triangle_p$. However, our intention is to exploit coordinate transformation properties to convert the integrals into equivalent expressions that will be precomputed on a fixed reference element. Here, the reference element will be denoted by $\triangle_r$.

3.5.1.1 The Mass Matrix

Assume that ɜ is a scalar function defined over the element domain $\triangle_r$. Also assume that is $\{x_s\}$ the set of points using which a scalar *Lagrangian* basis has been developed over $\triangle_r$. The mass

matrix T is generally defined as (3.50):

$$\mathsf{T}_{uv} = \int_{\triangle_p} \alpha_u \mathfrak{d}\, \alpha_v dx^3. \tag{3.50}$$

The mass matrix, can be used for evaluation of the second term on the LHS and the first term on the RHS of (3.45). We first pull the integration domain back into the reference element $\triangle_r$.

$$\mathsf{T}_{uv} = \int_{\triangle_r} \alpha_u \,\mathfrak{d}\, \alpha_v \det \mathsf{J}\xi^3 \tag{3.51}$$

$$\boxed{\varsigma \triangleq \det \mathsf{J}\, \mathfrak{d}} \tag{3.52}$$

Then we approximate (i.e. interpolate) the element-metric dependent term ς using the $\{\ell_s\}$ polynomial basis and sample ς at the interpolation point set $\{x_s\}$.

$$\boxed{\varsigma_s \triangleq \varsigma|_{x=x_s}} \tag{3.53}$$

Hence, T can be approximated as follows:

$$\mathsf{T}_{uv} \approx \sum_{0\leq s<\dim\{\ell_s\}} \varsigma_s \int_{\triangle_r} \alpha_u \alpha_v \ell_s \, d\xi^3 \tag{3.54}$$

If we define the fixed (precomputable) array $\overset{\circ}{\mathsf{T}}{}^{uvs}$ as (3.55), the final form of T can be written as (3.56).

$$\boxed{\overset{\circ}{\mathsf{T}}{}^{uvs} \triangleq \int_{\triangle_r} \alpha_u \alpha_v \ell_s \, d\xi^3} \tag{3.55}$$

$$\boxed{\mathsf{T}_{uv} \approx \sum_{0 \le s < \dim\{\ell_s\}} \varsigma_s \overset{\circ}{\mathsf{T}}^{uvs}} \tag{3.56}$$

In general, evaluation of (3.56), requires $\dim\{\ell_s\}$ floating point mul-add operations per matrix entry for the mass matrix T defined in (3.50).

3.5.1.2 The Stiffness Matrix

Here we assume $\mathfrak{u}$ to be a 3×3 tensor. The stiffness matrix entries S_{uv} are basically of the following form:

$$\mathsf{S}_{uv} = \int_{\triangle_p} \nabla_x \alpha_u \cdot \mathfrak{u} \cdot \nabla_x \alpha_v dx^3 \tag{3.57}$$

The stiffness matrix defined in (3.57), can be used for evaluation of the first term on the LHS (3.45). Transforming the integration domain from $\triangle_p$ into $\triangle_r$ one gets:

$$\mathsf{S}_{uv} = \int_{\triangle_r} (\mathsf{K}^\mathsf{T} \nabla_\xi \alpha_u)^\mathsf{T} \times \mathfrak{u} \times (\mathsf{K}^\mathsf{T} \nabla_\xi \alpha_v) \det \mathsf{J} d\xi^3 \tag{3.58}$$

$$\mathsf{S}_{uv} = \int_{\triangle_r} \nabla_\xi \alpha_u^\mathsf{T} \left(\mathsf{K} \times \mathfrak{u} \times \mathsf{K}^\mathsf{T} \right) \nabla_\xi \alpha_v \det \mathsf{J} d\xi^3 \tag{3.59}$$

We first define ς as the metric-dependent part of the integral:

$$\boxed{\varsigma \triangleq \det \mathsf{J}\, \mathsf{K} \times \mathfrak{u} \times \mathsf{K}^\mathsf{T}} \tag{3.60}$$

And take its' samples at the interpolation set $\{x_s\}$ to yield:

$$\boxed{\varsigma_{ijs} \triangleq \left[\left(\det \mathsf{J}\, \mathsf{K} \times \mathfrak{U} \times \mathsf{K}^{\mathsf{T}}\right)\right]_{ij}|_{x=xs}} \tag{3.61}$$

Hence, S_{uv} can be approximated as (3.62).

$$\mathsf{S}_{uv} \approx \sum_{i,j,0\leq s<\dim\{\ell_s\}} \varsigma_{ijs} \int_{\triangle_r} \frac{\partial \alpha_u}{\partial \xi_i} \frac{\partial \alpha_v}{\partial \xi_j} \ell_s d\xi^3 \tag{3.62}$$

For the sake of convenience, $\varkappa_{iujv}$ is defined as (3.63). Note that $\varkappa_{iujv}$ is symmetric in the sense of (3.64):

$$\boxed{\varkappa_{iujv} \triangleq \frac{\partial \alpha_u}{\partial \xi_i} \frac{\partial \alpha_v}{\partial \xi_j}} \tag{3.63}$$

$$\varkappa_{iujv} = \varkappa_{juiv} \tag{3.64}$$

Now, to add more symmetry to the expressions, we decide to compute $\frac{1}{2}(\mathsf{S}_{uv} \pm \mathsf{S}_{vu})$ instead of evaluating S_{uv} and S_{uv} separately. Thus, we get:

$$\frac{1}{2}(\mathsf{S}_{uv} \pm \mathsf{S}_{vu}) \approx \sum_{i,j,0\leq s<\dim\{\ell_s\}} \varsigma_{ijs} \int_{\triangle_r} \frac{\varkappa_{iujv} \pm \varkappa_{juiv}}{2} \ell_s d\xi^3 \tag{3.65}$$

Since the above expression is symmetric in $i \leftrightarrow j$, we can rewrite the summation in a computationally more efficient way , i.e. (3.66).

$$\frac{1}{2}(\mathsf{S}_{uv} \pm \mathsf{S}_{vu}) \approx \sum_{i\leq j,0\leq s<\dim\{\ell_s\}} (\varsigma_{ijs} \pm \varsigma_{jis}) \int_{\triangle_r} \frac{\varkappa_{iujv} \pm \varkappa_{juiv}}{2+2\delta_{ij}} \ell_s d\xi^3 \tag{3.66}$$

$$\boxed{\overset{\circ}{\mathsf{S}}{}^{\pm iujvs} \triangleq \int_{\triangle_r} \frac{\varkappa_{iujv} \pm \varkappa_{juiv}}{2+2\delta_{ij}} \ell_s d\xi^3} \tag{3.67}$$

$$\boxed{\mathsf{C}^{\pm}_{ijs} \triangleq \varsigma_{ijs} \pm \varsigma_{jis}} \tag{3.68}$$

Defining $\overset{\circ}{\mathsf{S}}{}^{\pm iujvs}$ and $\mathsf{C}^{\pm}_{ijs}$ as in (3.67) and (3.68), we arrive at the final form for evaluation of $\frac{1}{2}(\mathsf{S}_{uv} \pm \mathsf{S}_{vu})$ from which S_{uv} and S_{vu} can be obtained by simple addition.

$$\boxed{\frac{1}{2}(\mathsf{S}_{uv} \pm \mathsf{S}_{vu}) \approx \sum_{i\leq j, 0\leq s<\dim\{\ell_s\}} \mathsf{C}^{\pm}_{ijs} \overset{\circ}{\mathsf{S}}{}^{iujvs}} \tag{3.69}$$

The above formulation amounts to $6\dim\{\ell_s\}$ floating point mul-add operations per matrix entry. Noting that $\mathsf{C}^{-}_{iis} = 0$ further reduces the number of mul-add operations to $\frac{9}{2}\dim\{\ell_s\}$ per matrix entry. For each (u, v), actual S_{uv} and S_{vu} can be found be direct addition and subtraction of the evaluated sum/difference. This only adds an overhead of one floating point addition per matrix entry. Note that the $\overset{\circ}{\mathsf{S}}{}^{\pm iujvs}$ array does not depend to any element metric-specific properties. In other words, this is a fixed array which should be precomputed and stored for "look-up" operations.

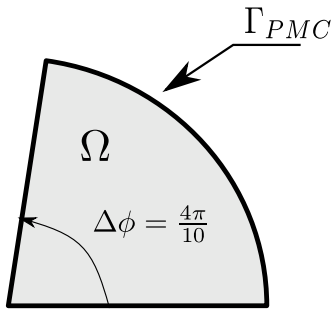
3.5.1.3 A Demonstrative Example

In order to have the proposed formulation tested, we shall solve for the eigenvalues of a scalar Helmholtz problem (3.70) , (3.71), for which the solution domain is a single element in the shape of a circular sector, see figure 3.2. Since the eigenvalues of this problem are analytically available, it is easy to have a check on the convergence of eigenvalues with respect to the polynomial order of the element basis $\{\alpha_i\}$ and the auxiliary Lagrangian basis $\{\ell_i\}$. Analytically, it is known that the dominant k for this problem equals to the first root of the first Bessel function $J_1(x)$, which equals to 3.83170597020751 [78]. Hence, with respect to the polynomial orders of the

element functional and auxiliary basis, the convergence curves of figure 3.3 and figure 3.4 are obtained. . Obviously Figure 3.4 exhibits the $h^{p_{basis}+1}$ behavior provided that the order of the auxiliary basis is high enough to capture the curvature up the the same order. This conforms to the standard results for *isoparametric* elements and confirms the reliability of the proposed method for evaluation of FE matrices [79].

$$\nabla \cdot \nabla\psi - k^2\psi = 0 \text{ in } \Omega \tag{3.70}$$

$$\nabla\psi \cdot \hat{n} = 0 \text{ on } \Gamma_{PMC} \tag{3.71}$$



(a)

Figure 3.2: The circular sector resonant cavity.

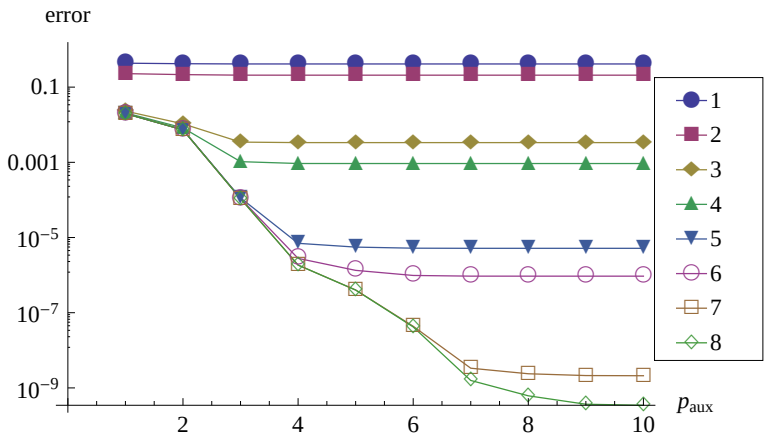


Figure 3.3: Dominant eigenvalue error as a function of auxiliary basis polynomial order p_{aux}. Various curves correspond to various polynomial orders of element functional basis p_{basis}.

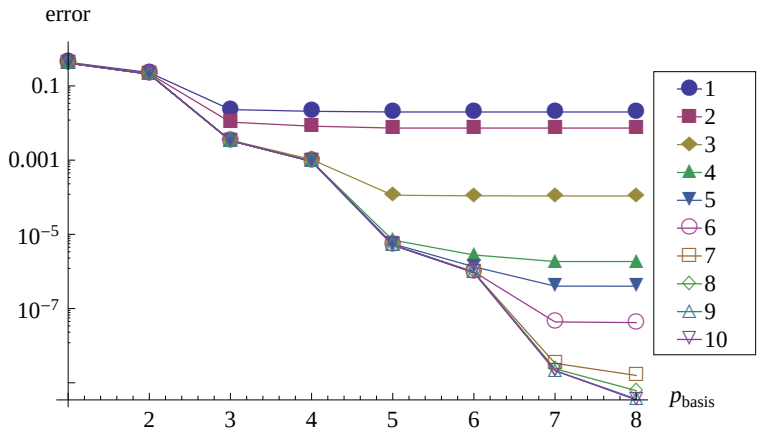


Figure 3.4: Dominant eigenvalue error as a function of functional basis polynomial order p_{basis}. Various curves correspond to various polynomial orders of auxiliary basis p_{aux}.

3.6 Discretization of I_{inf}

The rectangular topology of IEs translates into a tensor product nature of their basis functions and element matrices. This simplifies the construction of the basis functions and element matrices as the tensor product of single dimensional basis functions and FE matrices. Hence, for the required H^1 polynomial basis functions to be composed out of products of shape functions independently defined over *one dimensional/dimensions (1D)* intervals $\zeta \in [0,1]$ and $\xi_2 \in [0,1]$. Referring to (3.13), it is clear that the ϕ domain of each element directly determines b as $b = \phi_2 - \phi_1 = \Delta\phi$. Thus, the only metric dependent property of the element that is reflected in the element matrix is $\Delta\phi$. As depicted in figure 3.5, the IEs can be constructed by application of (3.13) onto a fixed master element. Let us assume that a tensor product Lagrangian basis $\{\alpha_{i_1}(\zeta)|0 \leq i_1 \leq O_\zeta\} \times \{\beta_{i_2}(\xi_2)|0 \leq i_2 \leq O_{\xi_2}\}$ is used to approximate S and T over the reference element $\square$.

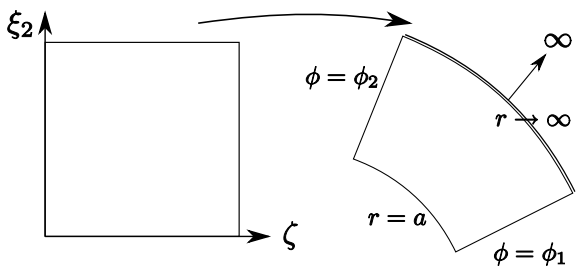


Figure 3.5: A rectangular master element $\square = \{(\zeta, \xi_2)|0 \leq \zeta, \xi_2 \leq 1\}$ (finite) is transformed to an infinite element.

$$\begin{cases} S = \sum_{i_1 i_2} S_{i_1 i_2} \alpha_{i_1}(\zeta)\beta_{i_2}(\xi_2) \\ T = \sum_{j_1 j_2} T_{j_1 j_2} \alpha_{j_1}(\zeta)\beta_{j_2}(\xi_2) \end{cases} \tag{3.72}$$

Thus, for an individual IE, $\lim_{R\to\infty} I_{inf,e}$ can be approximated as (3.73) where the $_e$ subscript stands for the IE of interest.

$$\sum_{i_1 i_2, j_1 j_2} \bar{T}_{j_1 j_2} S_{i_1 i_2} \Bigg(\Bigg(\Delta\phi \int_{\zeta=0}^{1} (\zeta^2 \frac{\partial \bar{\alpha}_{j_1}}{\partial \zeta} \frac{\partial \alpha_{i_1}}{\partial \zeta} + \alpha(\alpha-1)\bar{\alpha}_{j_1}\alpha_{i_1} - 2\jmath ka \frac{\partial \alpha_{i_1}}{\partial \zeta}\bar{\alpha}_{j_1}) d\zeta \Bigg) \Bigg(\int_{\xi_2=0}^{1} \bar{\beta}_{j_2}\beta_{i_2} d\xi_2 \Bigg) + \Bigg(\frac{1}{\Delta\phi} \int_{\zeta=0}^{1} \bar{\alpha}_{j_1}\alpha_{i_1} d\zeta \Bigg) \Bigg(\int_{\xi_2=0}^{1} \frac{\partial \bar{\beta}_{j_2}}{\partial \xi_2} \frac{\partial \beta_{i_2}}{\partial \xi_2} d\xi_2 \Bigg) + (\jmath ka + \alpha)\Delta\phi \bar{\alpha}_{j_1}\alpha_{i_1}|_{\zeta=1} \int_{\xi_2=0}^{1} \bar{\beta}_{j_2}\beta_{i_2} d\xi_2 \Bigg) = 0 \tag{3.73}$$

With the assumption that the polynomial expansions are completely real, (3.73) can be split into real and imaginary parts.

$$\mathfrak{Re}(\lim_{R\to\infty} I_{inf,e}) : \sum_{i_1 i_2, j_1 j_2} \bar{T}_{j_1 j_2} S_{i_1 i_2} \Bigg(\frac{1}{\Delta\phi} \Bigg(\int_{\zeta=0}^{1} \bar{\alpha}_{j_1}\alpha_{i_1} d\zeta \Bigg) \Bigg(\int_{\xi_2=0}^{1} \frac{\partial \bar{\beta}_{j_2}}{\partial \xi_2} \frac{\partial \beta_{i_2}}{\partial \xi_2} d\xi_2 \Bigg) + \Delta\phi \Bigg(\int_{\zeta=0}^{1} (\zeta^2 \frac{\partial \bar{\alpha}_{j_1}}{\partial \zeta} \frac{\partial \alpha_{i_1}}{\partial \zeta} + \alpha(\alpha-1)\bar{\alpha}_{j_1}\alpha_{i_1}) d\zeta + \alpha\, \bar{\alpha}_{j_1}\alpha_{i_1}|_{\zeta=1} \Bigg) \Bigg(\int_{\xi_2=0}^{1} \bar{\beta}_{j_2}\beta_{i_2} d\xi_2 \Bigg) \Bigg) = 0 \tag{3.74}$$

$$\mathfrak{Im}(\lim_{R\to\infty} I_{inf,e}) : \sum_{i_1 i_2, j_1 j_2} \bar{T}_{j_1 j_2} S_{i_1 i_2} ka\Delta\phi \Bigg(\bar{\alpha}_{j_1}\alpha_{i_1}|_{\zeta=1} - 2\int_{\zeta=0}^{1} \frac{\partial \alpha_{i_1}}{\partial \zeta} \bar{\alpha}_{j_1} d\zeta \Bigg) \Bigg(\int_{\xi_2=0}^{1} \bar{\beta}_{j_2}\beta_{i_2} d\xi_2 \Bigg) = 0 \tag{3.75}$$

In order to develop a matrix formulation for the $\lim_{R\to\infty} I_{inf,e}$ one needs to convert the double indices of say $S_{i_1 i_2}$ into a single index format. There are many ways for construction of such equivalent indices. Considering that the total dimension of the element polynomial space equals to $(1 + O_\zeta)(1 + O_{\xi_2})$, the following formulation will be used for construction of the required equivalent indices.

$$i(i_1, i_2) = i_1 + (i_2 - 1)(1 + O_\zeta) \tag{3.76}$$

If the equivalent single indices are used in (3.74) and (3.75), $\lim_{R\to\infty} I_{inf,e}$ can be formulated as it follows in (3.77) where $\mathsf{S} = [\mathsf{S}_i] = [S_{i_1 i_2}]$ and $\mathsf{T} = [\mathsf{T}_i] = [T_{i_1 i_2}]$ respectively represent row vectors containing the sampled values of S and T functions and where the S_{inf} matrix is defined according to (3.78). In analogy to its FE counterpart, $\mathsf{S}_{\mathsf{inf}}$ will be regarded as the IE *stiffness* matrix. The $\overset{\circ}{\mathsf{I}}_{r1}$, $\overset{\circ}{\mathsf{I}}_{r2}$ and $\overset{\circ}{\mathsf{I}}_i$ matrices introduced in (3.78) are constant matrices who's definitions follow in (3.79), (3.80) and (3.81). The S_{inf} matrix evaluated here corresponds to a single IE and the global matrix must be assembled from the complete ensemble of such individual IE matrices.

$$\lim_{R\to\infty} I_{inf,e} : \bar{\mathsf{T}}^{\mathsf{T}} \times \mathsf{S}_{\mathsf{inf}} \times \mathsf{S} = 0 \tag{3.77}$$

$$\mathsf{S}_{\mathsf{inf}} = \frac{1}{\Delta\phi}\overset{\circ}{\mathsf{I}}_{r1} + \Delta\phi\,\overset{\circ}{\mathsf{I}}_{r2} + \jmath ka\Delta\phi\,\overset{\circ}{\mathsf{I}}_i \tag{3.78}$$

$$\overset{\circ}{\mathsf{I}}_{r1} = [\overset{\circ}{\mathsf{I}}_{r1ij}] = \left(\int_{\zeta=0}^{1} \alpha_{i_1}\bar{\alpha}_{j_1} d\zeta\right)\left(\int_{\xi_2=0}^{1} \frac{\partial\beta_{i_2}}{\partial\xi_2}\frac{\partial\bar{\beta}_{j_2}}{\partial\xi_2} d\xi_2\right) \tag{3.79}$$

$$\overset{\circ}{\mathsf{I}}_{r2} = [\overset{\circ}{\mathsf{I}}_{r2ij}] = \left(\int_{\zeta=0}^{1} (\zeta^2\frac{\partial\alpha_{i_1}}{\partial\zeta}\frac{\partial\bar{\alpha}_{j_1}}{\partial\zeta} + \alpha(\alpha-1)\alpha_{i_1}\bar{\alpha}_{j_1})d\zeta + \alpha\alpha_{i_1}\bar{\alpha}_{j_1}|_{\zeta=1}\right)\left(\int_{\xi_2=0}^{1} \beta_{i_2}\bar{\beta}_{j_2} d\xi_2\right) \tag{3.80}$$

$$\overset{\circ}{\mathsf{I}}_i = [\overset{\circ}{\mathsf{I}}_{iij}] = \left(\alpha_{i_1}\bar{\alpha}_{j_1}|_{\zeta=1} - 2\int_{\zeta=0}^{1} \frac{\partial\alpha_{i_1}}{\partial\zeta}\bar{\alpha}_{j_1} d\zeta\right)\left(\int_{\xi_2=0}^{1} \beta_{i_2}\bar{\beta}_{j_2} d\xi_2\right) \tag{3.81}$$

3.7 Results

3.7.1 Rectangular Scatterer

The model problem solved here is a PEC rectangular scatterer as depicted in figure 3.6. The model is excited by an incoming plane wave incident from $\hat{x} + \hat{y}$ direction. The side of the scatterer measures 4λ where λ represents the free space wavelength. Figure 3.7, plots the scattered field F/IE solution to the problem obtained using 7^{th} order elements.

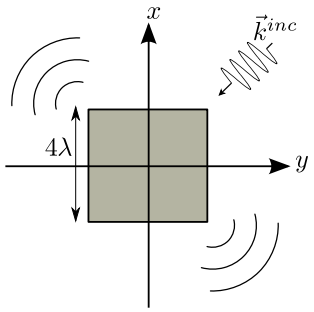


Figure 3.6: Problem model for the rigid (PEC) scatterer.

3.7.2 Effect Of Nodal Sets On F/IE Matrix Condition Numbers

A well know issue with the use of IEs is the growth of their matrix condition numbers with problem size. The tensorial nature of these elements has been exploited for development of suitable preconditioners for F/IE problems [42–44, 80]. Another techniques that can significantly improve the condition number of IEs matrices, is based on the choice of basis functions. In the context of spectral FEs, this is achieved by means of manipulating and optimizing the location of interpolation nodal set. Chapter 4 is dedicated to this topic. However, here the topic is brought up in the context of IEs. Figure 3.8 plots the condition number of the assembled F/IE matrix for the problem of figure 3.6. A detailed description of the nodal sets mentioned in figure 3.8 can be found in chapter 4. The particular discretization used for this example involved 1234 triangles and 90 infinite elements. Unlike the studies of chapter 4, some of the choices of the nodal sets have resulted in almost identical curves. This is because, the condition number of the assembled matrix

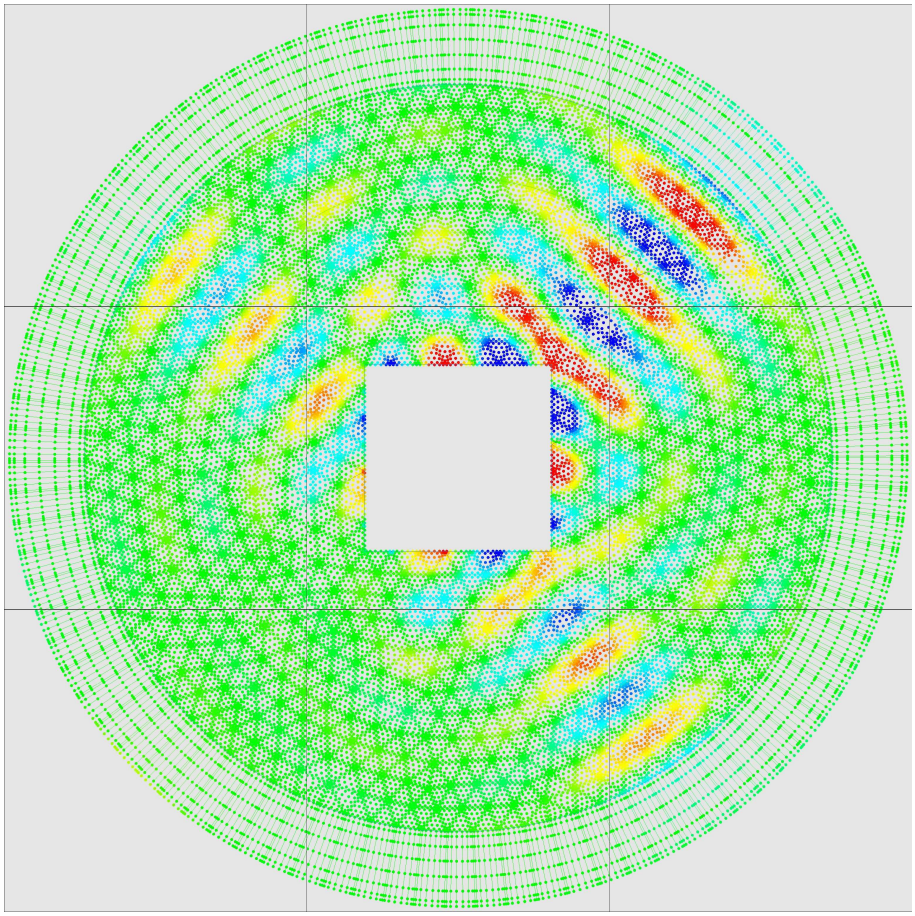


Figure 3.7: FEM solution for the scattered field from rigid rectangular scatterer. The outer circular rim represent the infinite element region.

is dominated by the IE matrix for which the interpolation nodal set follows a tensor product of the boundary nodes of triangular FEs. Hence, any two nodal sets that have identical patterns on triangle edges, will exhibit similar behavior with respect to IE matrix condition numbers. As evident from figure 3.8, at least for higher order elements, the condition numbers of F/IE matrices can be significantly reduced by choice of improved interpolation nodal sets. In this example, no other IE preconditioning techniques has been applied so that the resulting improvements can be observed purely as a function of the choice of nodal set. Nevertheless, an efficient implementation of the F/IE technique should be adopted by integration of efficient preconditioners [42–44, 80], such that the combined effects of the preconditioners and nodal sets can be exploited. The combination of these methods, however, can be the topic of future research.

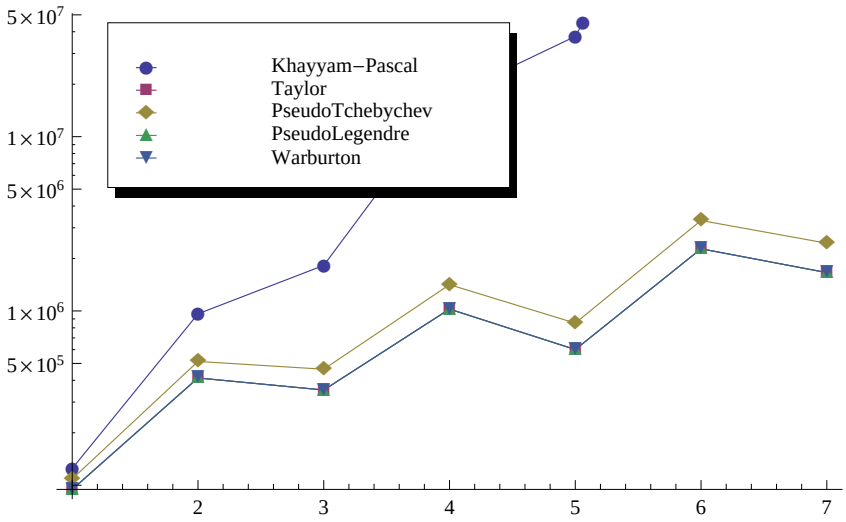


Figure 3.8: Condition number of the F/IE matrix as a function of polynomial order for various choices of interpolation nodal sets.

Chapter 4

Nodal Sets and FE Matrix Condition Numbers

4.1 Introduction

Spectral FEMs rely on spectral interpolation of functions over the triangulated problem domain. It is a known fact that the spatial distribution of the interpolation points play critical role in quality of the resulting interpolation operators. A variety of methods have been adapted for the construction of optimal precision sets (interpolation nodes) in d-simplices. The interpolatory quality of such precision sets is technically gauged using the associated *Lebesgue* constant. With few exceptions, most of the related literature are, in this regard, concerned with interpolation-theoretical aspects of such improved interpolation nodal sets. This is while from FE analysis point of view, we are mainly concerned with achieving better matrix condition numbers. Our experiments show that better FE matrix conditions do not necessarily translate to better interpolatory properties of the involved nodal sets and vice versa. This is in spite of the fact that the concrete foundation of Lebesgue's lemma often leads to the impression that nodes with smaller Lebesgue constants give better results when applied in FE analysis. This chapter, however, examines some available precision sets for their influence on FE matrix condition numbers. It is hence observed that, in

H^1 (*Laplace* 's, Poisson's and Helmholtz problems), Fekete nodes offer better condition numbers although they do not have the best Lebesgue constants. For $H(\nabla\wedge)$ elements (problems in magnetostatics and EM), however, the situation is more complex. At the zero frequency end, the stiffness matrix becomes underdetermined and the resulting matrix problem can be solved only by imposition of extra constraints. Thus, the quality of the interpolation nodal set must be judged with respect to the condition numbers of both FE matrices and constraint equations. In this case, *warp & blend* nodes exhibit more robust performance.

4.1.1 Background

This section is dedicated to the development of the required common language and conventions used throughout the article. It is thus composed in the briefest possible way.

In the context of approximation theory, the quality of an interpolation operator is often gauged by the associated Lebesgue constant. Assuming that the interpolation operator $\mathcal{L}$ is defined over a domain of interest Ω, Lebesgue constant Λ is defined as $\Lambda = \sup_{x\neq 0} \frac{||\mathcal{L}x||}{||x||}$. The merit of Λ is better understood when, according to Lebesgue's lemma, it is treated as a gauge for the quality of the interpolation operator $\mathcal{L}$.

$$||f - \mathcal{L}f|| \leq (1 + \Lambda)||f - h|| \tag{4.1}$$

In (4.1), h represents the best possible polynomial interpolation to f in the $||\bullet||$ norm [81]. Furthermore, with $||\bullet|| = ||\bullet||_\infty$, it can be shown that $\Lambda = \max_x \sum_i |l_i(x)|$ where $l_i(x)$ are the Lagrangians over the interpolation domain. In general, the infinity norm is what we mean when speaking about Λ. Bloom et. al. [82] show that the following condition is a must for the spectral convergence of an interpolation operator.

$$\lim_{p\to\infty} \Lambda^p = 1 \tag{4.2}$$

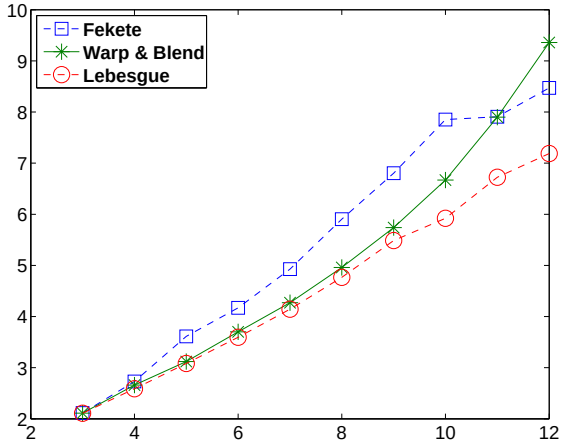


Figure 4.1: Lebesgue's Constant versus polynomial order.

Figure 4.1, compares Lebesgue constants for some families of precision sets. In figure 4.1, it is apparent that warp & blend nodes have better Lebesgue constants for polynomial orders up to 11. Hence, if the quality of an interpolation operator (from FEM matrix condition's perspective) could be judge solely based on the associated Lebesgue constant, then warp & blend nodes are expected to yield best FEM matrix condition numbers.

Direct minimization of Lebesgue's constant is a cumbersome global optimization problem which was only recently attempted by Roth [83]. Prior to Roth, constructing optimal precision sets in a hope for achieving better Λ has been attempted using a variety of alternative measures. For example, Chen and *Babuska* [19] construct their nodes by means of optimizing (4.3) while Hesthaven [22] uses minimization of an electrostatic energy for obtaining the so called *electrostatic* nodes.

$$\beta = \int_{\Omega=\triangle} \sum_i l_i^2 \, d\Omega \tag{4.3}$$

Taylor et. al, use a maximization of the *Generalized Vandermonde Determinant (GVD)* over triangular domains [24]. By definition, such nodes (that maximize the *determinant* of *Generalized Vandermonde Matrix (GVM)*) are regarded as Fekete nodes. Before the release of Roth's results on Lebesgue nodes, Fekete nodes were known to provide the lowest Lebesgue constants. Fekete nodes are known to coincide with *Gauss-Lobatto-Legendre (points) (GLL)* nodes over one dimensional and tensor product domains [84]. Furthermore, the Fekete criterion is known to bound

the Lebesgue function defined in (4.4) to N where N is the number of interpolation nodes.

$$\Lambda = \max_x \sum_i |l_i(x)| \leq N \tag{4.4}$$

In triangular and simplex shaped domains, while Fekete nodes do satisfy the convergence criterion (4.2), they do not necessarily lead to minimum Lebesgue constants. On the other hand, maximization of the Vandermonde determinant is not guaranteed by minimization of of Λ in (4.4). This point will contrasted as we proceed with our presentation in the following sections.

4.1.2 Interpolation Precision Sets

Despite the long history of spectral methods and in the presence of a dominant consensus on the advantages of hierarchic methods, certain areas in the field of interpolatory FEM are still open to exploration. Due to their interpolatory nature, spectral methods find intrinsic connections with approximation theory where Lebesgue's constant is widely used as a gauge for the quality of interpolation operators. The strong foundation of Lebesgue's lemma often leads to the impression that nodes with the best Lebesgue measure should yield the best performance for FE analysis. Our investigations in this article, however, indicate that lower Lebesgue constants do not necessarily lead to better FE matrices. For example, for problems posed in the H^1 space, precision sets with lower Lebesgue constants exist that fail to outperform Fekete [24] nodes in regards to FE matrix condition numbers. In this article, the following precision sets will be examined for their potential capabilities in improving FE matrices:

1. Warp & blend nodes [28]
2. Fekete nodes [24]
3. *Pseudo -Legendre* nodes: Introduced as "*Lobatto* interpolation grid" in [25, 27].
4. Pseudo-*Tchebychev* nodes: Similar to pseudo-Legendre [27] nodes except that *Gauss--Lobatto-Tchebychev (points) (GLT)* nodes have been used on the boundary.

5. *Khayyam-Pascal isometric (KPI)* nodes.

Higher order interpolatory $H(\nabla\wedge)$ basis were previously developed by Wilton et. al. [12]. Their construction of the basis is inherently based on (shifted) *Silvester* polynomials and isometric nodes [48, 73]. Thus, their approach is here modified and extended for arbitrary choices of precision sets. Indeed, by 'arbitrary' it is not meant that any random distribution of nodes is applicable to the development of a FE basis.

4.2 H^1 Basis

FE solution of the Poisson's equation involves in the solution of a linear system of algebraic equations $A\ x\ =\ Bf$, in which A and B are often regarded as stiffness and *mass* matrices respectively. In element-wise fashion, these matrices are evaluated according to (4.5) and (4.6) [70, 73].

$$A = [A_{ij}] = \left[\int_{\Omega=\triangle} \nabla\alpha_i.\nabla\alpha_j \; d\Omega\right] \tag{4.5}$$

$$B = [B_{ij}] = \left[\int_{\Omega=\triangle} \alpha_i \; \alpha_j \; d\Omega\right] \tag{4.6}$$

In (4.5) and (4.6), $\{\alpha_i\}$ are the so called element *shape* functions and coincide with the Lagrangians $\{l_i\}$ defined over the element. Condition numbers of stiffness and mass matrices play critical role in the accuracy and pace of convergence of the iterative process involving the solution of FE matrices. Hence, we would study the condition number of these matrices as a function of the choice of the precision sets. First, as presented in section 4.2.1, some analytical observations on the condition numbers of these matrices will be presented. These analytical results will be compared to the numerical results obtained for two cases: A single element problem and a real FE problem involving a complex triangular mesh.

4.2.1 Basic Bound Analysis

Let us consider mass matrix as defined in (4.6). First, the Lagrangians (or the shape functions) must be developed. For this purpose, one needs to construct the GVM. *Koornwinder* 's p-Order orthogonal basis [83, 85] defined in (4.7) will be used for the construction of the Lagrangians and the associated GVM.

$$\begin{aligned} &0 \leq i+j \leq p, 0 \leq i,j \ : \\ &k_{ij}(x,y) \triangleq \sqrt{2(2i+1)(i+j+1)} P_i^{(0,0)}(\tfrac{2x+y-1}{y-1})(1-y)^i P_j^{(2i+1,0)}(2y-1) \end{aligned} \tag{4.7}$$

The dimension of the polynomial space $P^p(\triangle)$ that is spanned by the Koornwinder basis is denoted by dm where $dm = \frac{(p+1)(p+2)}{2}$. Now, suppose that the Koornwinder basis components are linearly ordered and indexed by $\{k_i(x,y) | 1 \leq i \leq dm\}$ rather than by the double indexed $k_{ij}(x,y)$ of (4.7). With this assumption, the GVM can be developed as (4.8) in which x_js denote (x,y) location of various interpolation nodes over the triangle:

$$V = [V_{ij}] \triangleq [k_i(x_j)] \tag{4.8}$$

Furthermore, utilizing the so called delta property of the Lagrangians their associated coefficients matrix $L = [L_{ij}]$ is derived as in (4.9).

$$l_i(x) = \textstyle\sum_{i=1}^{dm} L_{ij} k_j(x) \ \therefore$$

$$L = [L_{ij}] = V^{-1} \rightarrow \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_{dm} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{dm} \end{bmatrix} = V^{-1} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_{dm} \end{bmatrix} \tag{4.9}$$

As a result, with some algebraic manipulation, one arrives at the following formulations for A and B [83]:

$$\begin{aligned} &A = V^{-1}\left[\int_{\Omega=\triangle} \nabla k_i.\nabla k_j \; d\Omega\right] V^{-1T} = V^{-1}C_A V^{-1T}, \\ &\text{where } C_A = \int_{\Omega=\triangle} \nabla k_i.\nabla k_i \; d\Omega \end{aligned} \tag{4.10}$$

$$\begin{aligned} &B = V^{-1}\left[\int_{\Omega=\triangle} k_i k_j \; d\Omega\right] V^{-1T} = V^{-1}C_B V^{-1T}, \\ &\text{where } C_B = \int_{\Omega=\triangle} k_i.k_i \; d\Omega = [\delta_{ij}] = I \end{aligned} \tag{4.11}$$

In general, the condition number of M, or $\kappa(M)$, is defined using (4.12) in which $||\bullet||$ denotes any appropriate (induced) matrix norm.

$$\kappa(M) = ||M||\;||M^{-1}|| \tag{4.12}$$

However, care must be taken in cases (as in this case) where C matrix is singular. The actual situation when solving FEM matrices is that the matrix singularities are avoided by introduction of extra constraints and/or confinement of the RHS to the *range-space* of the matrix operator [1, 58]. For this reason, the usual definition of induced matrix condition number, i.e. (4.12), must be replaced with (4.13) where M^+ is an appropriate *pseudo-inverse* of M such as the *Moore -Penrose* pseudo-inverse. For non-singular square matrices M^+ is identical to M^{-1} and thus we shall use (4.13) as a general definition.

$$\kappa(M) \triangleq ||M||\;||M^{+}|| \tag{4.13}$$

From numerical error analysis point of view (linear system solution), the definition of $\kappa(M)$ makes sense only when $||\bullet||$ is a vector induced norm. Application of other (non-induced norms) to (4.12) or (4.13) would be meaningless unless if the deployed norm is equivalent to some other vector induced norm. In our case, using *Frobenius* norm together with (4.10) and (4.11) , $\kappa(A)$

and $\kappa(B)$ can be further expanded as (4.14).

$$\kappa_F(M) = ||V^{-1}C_M V^{-1T}||_F \; ||V^T C_M^+ V||_F \leq \kappa_F(V)^2 \kappa_F(C_M) \tag{4.14}$$

In this development, the facts that Frobenius norm (denoted by the F subscript) is symmetric and *submultiplicative* are exploited. In (4.14), M is a general matrix representing A or B following the appropriate choice of C_M, i.e. C_A or C_B respectively. Here, it must be noted that the Frobenius norm is equivalent to the l^2 norm and thus the practical usefulness of (4.12) is not affected by the introduction of this norm. The equivalence is mathematically stated as (4.15).

$$||M||_2 \leq ||M||_F \leq \sqrt{n}\,||M||_2 \tag{4.15}$$

On the other hand, taking advantage of the norm equivalence (4.15), one concludes that :

$$M \in \mathbb{R}^{n\times n}, \quad \kappa_2(M) \leq \kappa_F(M) \leq n\,\kappa_2(M) \tag{4.16}$$

$$M \in \mathbb{R}^{n\times n}, \quad \frac{1}{n}\kappa_F(M) \leq \kappa_2(M) \leq \kappa_F(M) \tag{4.17}$$

In (4.16) and (4.17), the subscript 2 indicates the use of l^2 induced matrix norm and n indicates the dimension of the involved square matrices.

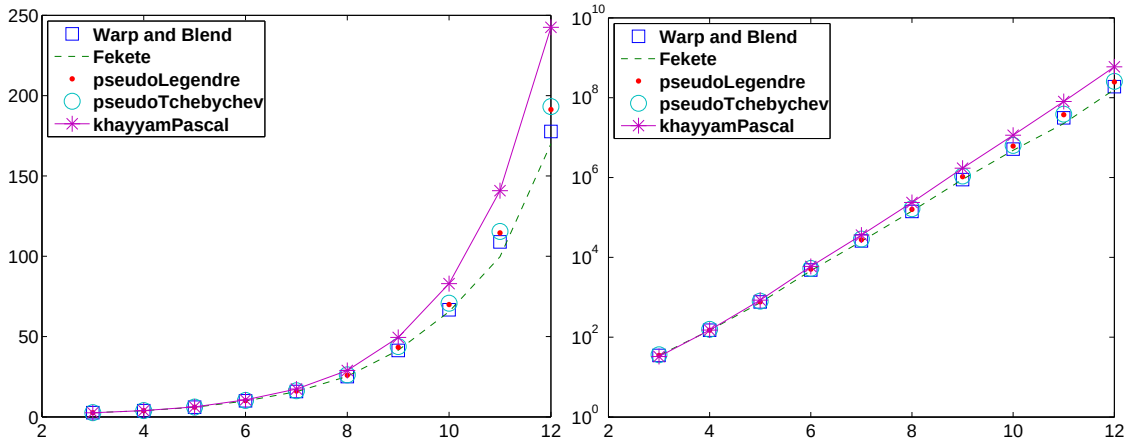
Now, substitution of (4.16) and (4.17) in (4.14) yields the following inequalities.

$$\frac{1}{n}\kappa_F(M) \leq \kappa_2(M) \leq \kappa_F(M) \leq \kappa_F(C_M)\kappa_F(V)^2 \quad \therefore \tag{4.18}$$

$$\kappa_2(M) \leq \kappa_F(M) \leq \kappa_F(V)^2 \leq n^2\kappa_F(C_M)\kappa_2(V)^2 \leq n^3\kappa_2(C_M)\kappa_2(V)^2 \tag{4.19}$$

It must be emphasized that M in (4.19) is a general matrix representing A or B following the appropriate choice of C_M, i.e. C_A or C_B. From (4.19), it is understood that bounding either

the Frobenius or the l^2 induced condition numbers of V bounds both Frobenius and l^2 induced condition numbers of M (A and B in particular). The argument, however, can be further extended to some other norms using existing norm equivalence relations. The main point here, is to take advantage of the symmetric and submultiplicative nature of the Frobenius norm as presented in the derivations above. Note that, in (4.19) V is the only entity that is dependent on the choice of the interpolation precision set. Figure 4.2, plots the condition number of GVM (as a function of polynomial order) for various precision sets.



(a) $\kappa^{\frac{e}{10}}$ versus polynomial order. (b) $log_e(\kappa)$ versus polynomial order

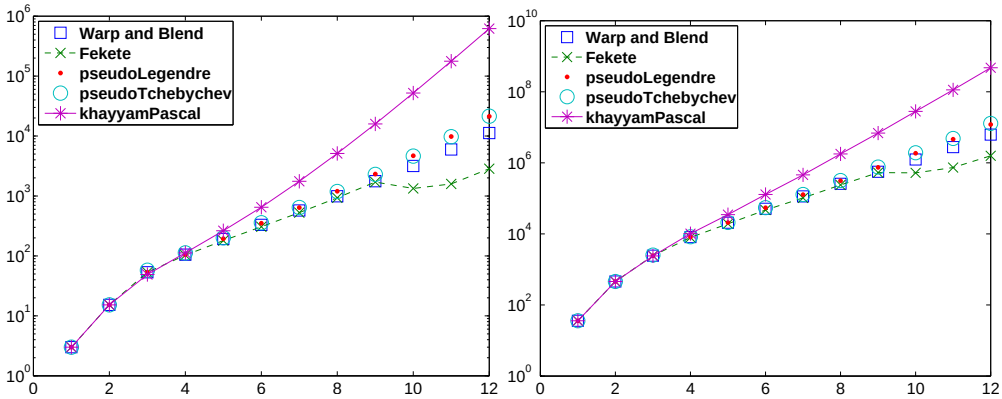
Figure 4.2: Condition number of the GVM (for H^1 elements) defined in (4.8) versus polynomial order on the right triangle $\triangle = \{(x, y)|0 \leq x, y, 0 \leq x + y \leq 1\}$.

4.2.2 The Single Element Case

We have examined various precision sets for their effect on the condition number of A and B matrices. The studied precision sets are the five sets mentioned earlier in section 4.1.2. Figure 4.3, depicts the condition number of single element stiffness and mass matrices for polynomial orders up to 12. From the figure, it is apparent that Fekete nodes generally provide better condition numbers for both stiffness and mass matrices. The fact that both A and B follow the same trend in figure 4.3 is in good agreement with the analysis of section 4.2.1.

At this point, it is necessary to disambiguate the definition of $\kappa(A)$, since it is basically a *rank deficient* matrix. The rank deficiency of A originates from the fact that the polynomial space

that is spanned by the basis functions $\{\alpha_i\}$ includes a proper (one-dimensional)*null-space* of the gradient operator ∇, i.e. a constant term. This one-dimensional null-space can be identified as the space of constant polynomials defined over the element. The existence of this proper null-space contributes to the rank deficiency of A (and equivalently C_A) as it can be seen from A's definition in (4.5).



(a) stiffness matrix (b) mass matrix

Figure 4.3: Condition number versus polynomial order of the stiffness and mass matrices respectively defined in (4.5) and (4.6) on the right triangle $\triangle = \{(x, y)|0 \leq x, y, 0 \leq x + y \leq 1\}$

4.2.3 The General Case

To further examine the effect of various precision sets on the performance of FE matrices, we set up a Poisson's problem over a square shaped domain $\square = \{(x, y)| - 2 \leq x, y \leq 2\}$. Suppose that the problem is formulated as (4.20). Furthermore, suppose that the boundary conditions and the distributed force terms (charge density in the language of *electrostatics*) are intentionally chosen in such a way that $u(x, y) = e^{-2(x^2+y^2)}$ satisfies (4.20) and its associated boundary conditions.

$$\nabla^2 u(x, y) = -\frac{\rho}{\epsilon}, \quad \text{boundary conditions}. \tag{4.20}$$

The abovementioned PDE problem is then triangulated and the associated global FE matrices A and B are assembled with different choices of precision sets. Figures 4.4 and 4.5, provide examples of such HO triangulations (FE mesh) using KPIs nodes and Fekete nodes. The resulting

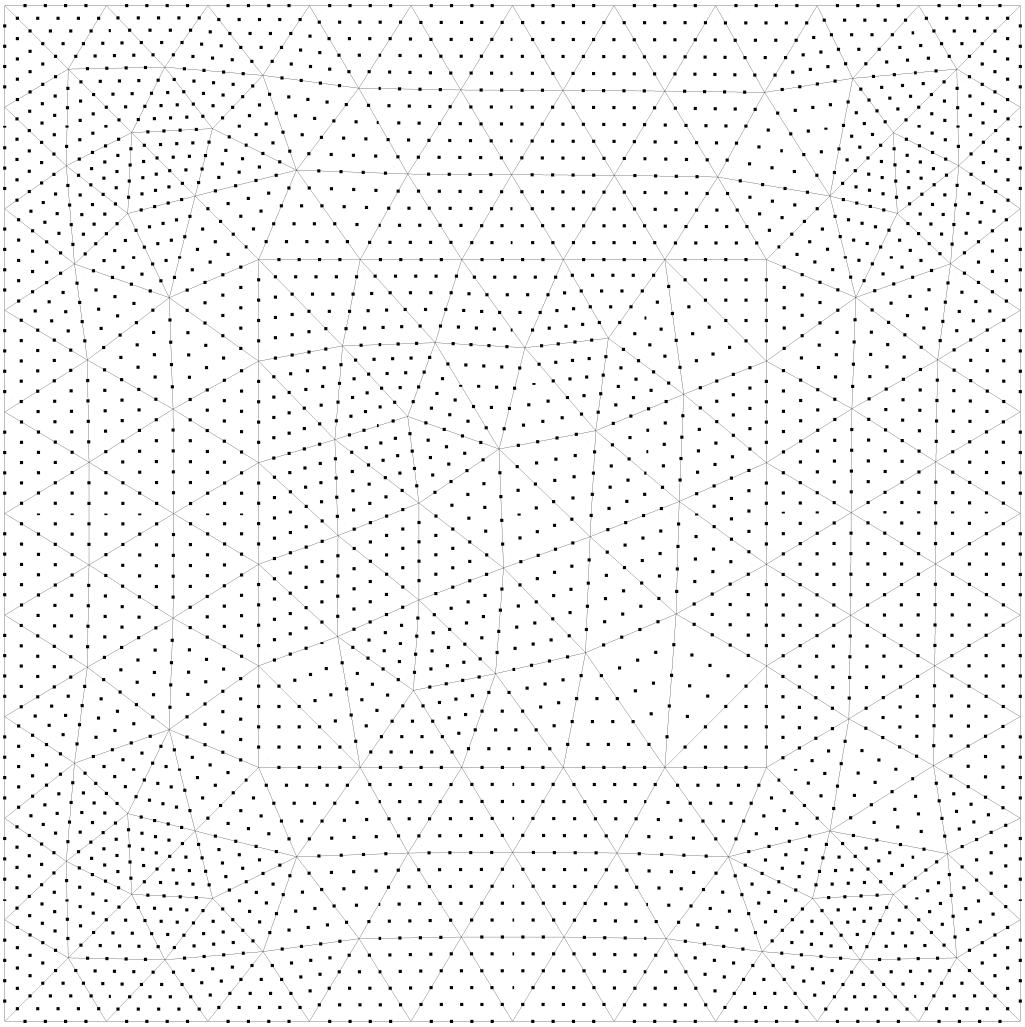


Figure 4.4: A 5^{th}-order FE mesh over the rectangular domain $\square$ comprising of 258 elements and 3326 nodes with Khayyam-Pascal Isometric Nodes.

matrix problems are then solved using an iterative *generalized minimal residual (GMRES)* solver. Figures 4.6, 4.7 and 4.8, depict the convergence history of the iterative solver for various choices of precision sets and various polynomial orders. Practically speaking, the convergence rate of the matrix solution is determined by the slope of these curves. As apparent from the mentioned plots, for all polynomial orders up to 12, Fekete nodes result in superior FE matrices and end up with subsequent improvements in the convergence rates of the associated matrix problems. This is in good agreement with the predictions made in section 4.2.2. The superiority of Fekete nodes is more contrasted at higher polynomial orders, specially when $p = \{10, 11, 12\}$. Similar contrasts are observable in the condition numbers of single element FE matrices plotted in figure 4.3.

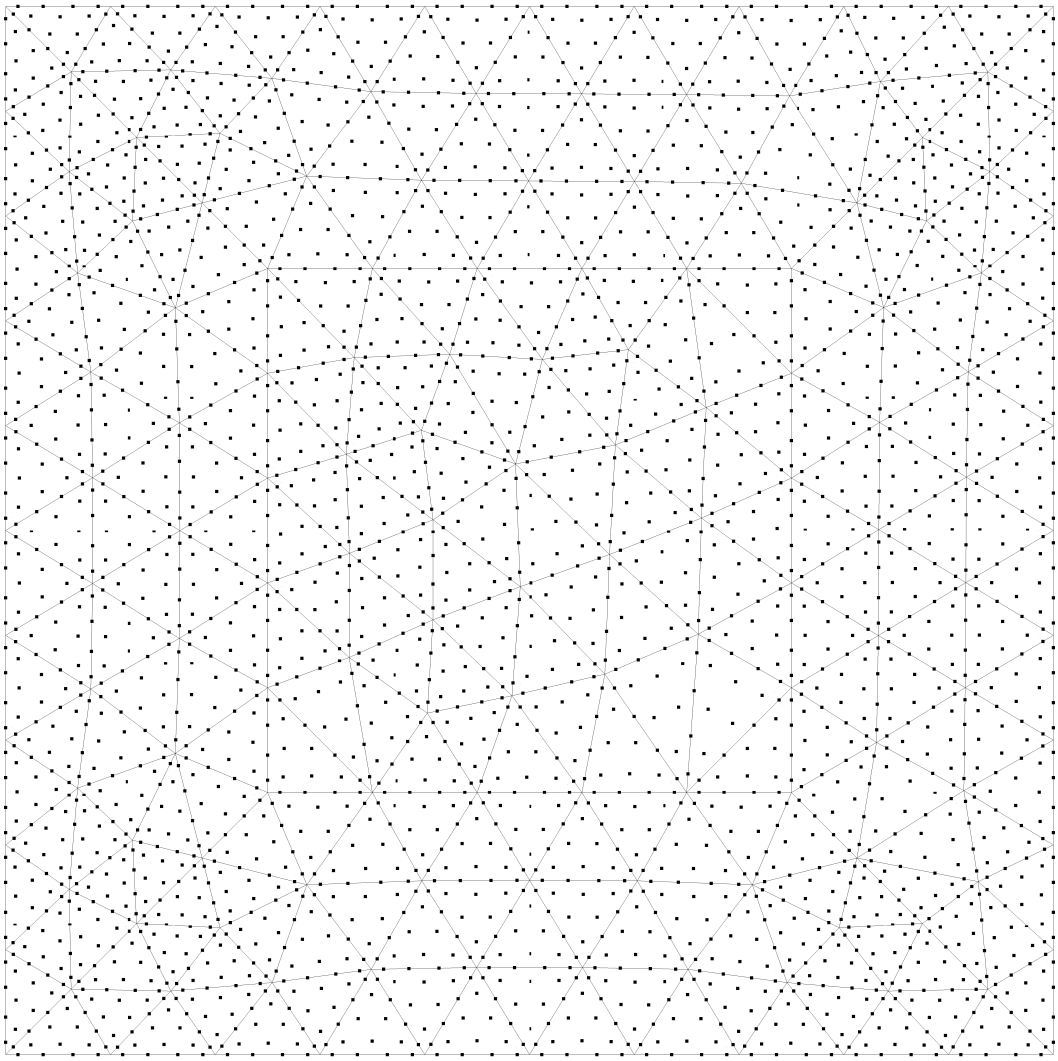
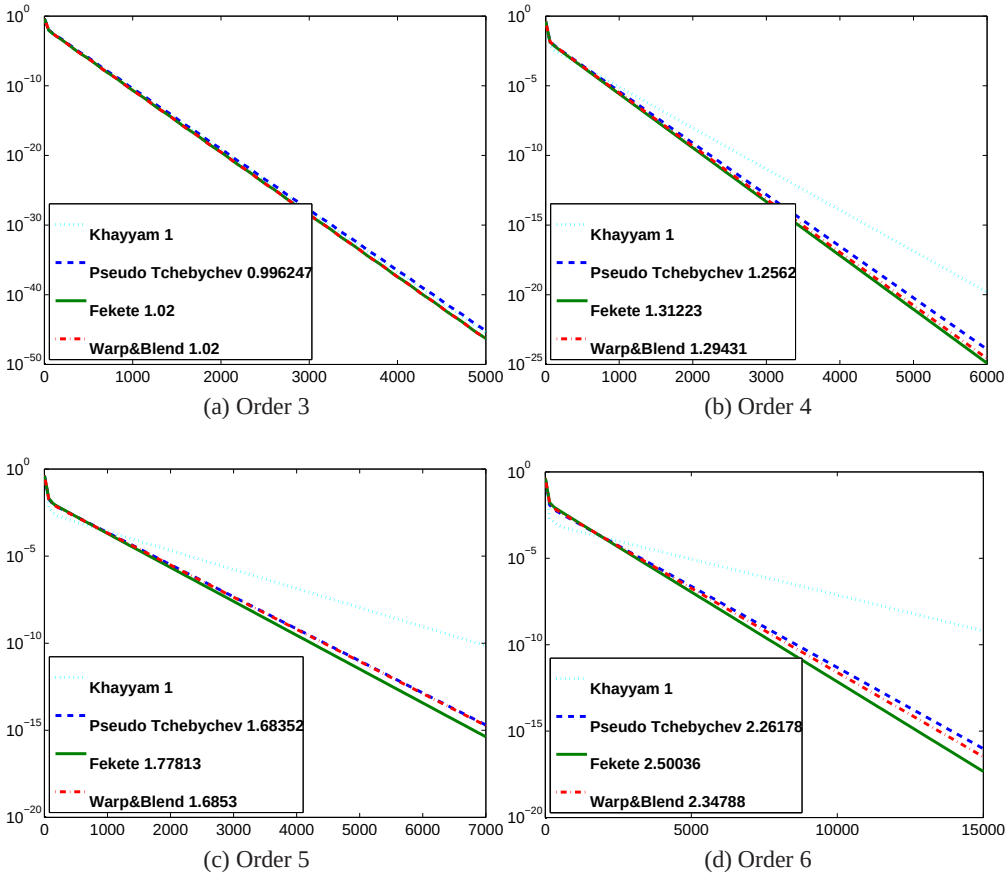


Figure 4.5: A 5^{th}-order FE mesh over the rectangular domain $\square$ comprising of 258 elements and 3326 nodes with Fekete Nodes.

4.2.4 H^1 Conclusion

Through the experiments conducted in section 4.2.2 and section 4.2.3, it is observed that (among the studied precision sets) Fekete nodes yield the best performance (among the examined sets) for H^1 problems such as Poisson's and Laplace's equations. This can be explained by referring back to (4.19). From (4.19), it is understood that the condition number of the GVM plays critical role in bounding the condition numbers of A and B. Furthermore, Fekete nodes are defined to be the configuration of nodal sets that would maximize the determinant of the GVM. This means that using Fekete nodes, the row vectors associated to the GVM are kept as orthogonal as possible to each other so that the determinant of the GVM (or in other words the *hypervolume*



(a) Order 3 (b) Order 4

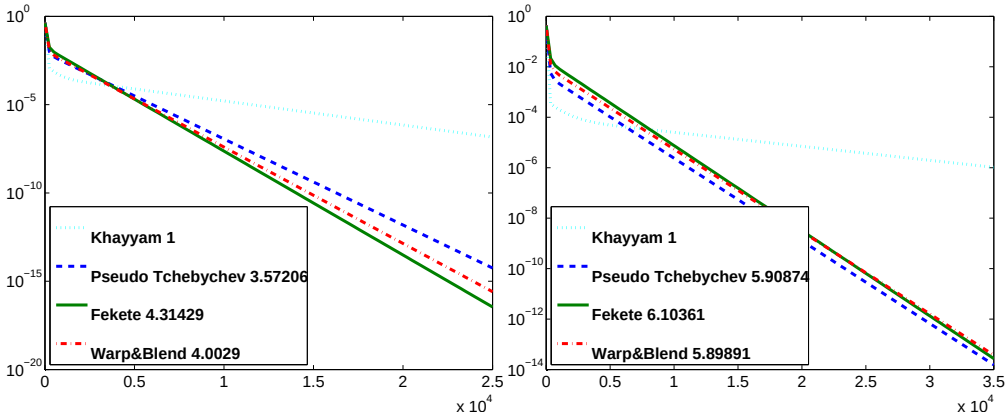
(c) Order 5 (d) Order 6

Figure 4.6: GMRES matrix solver's relative error as a function of iteration number. See how the choice of basis affects the convergence rate. The numbers in the legend are the slope of the curves relative to that of *Khayyam-Pascal* isometric nodes. See more plots in figures 4.7 and 4.8.

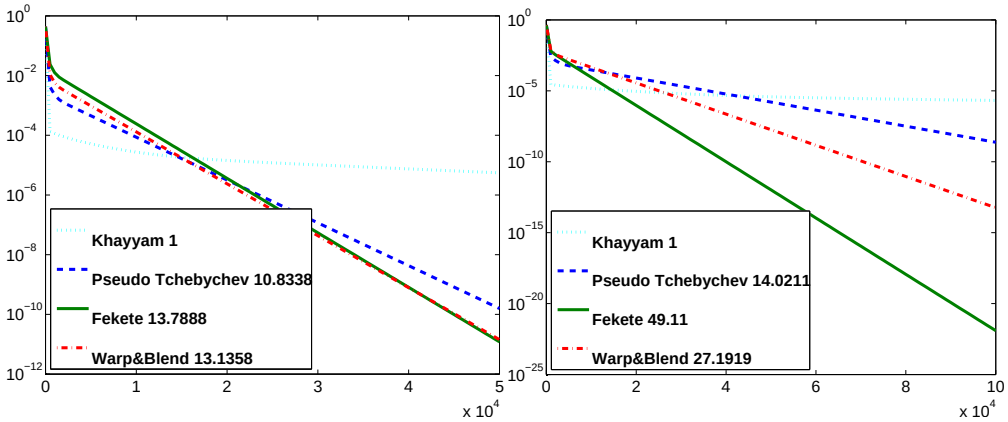
of the hypercube spanned by the row vectors of the GVM) is maximized. On the other hand, orthogonalization of the row vectors in GVM directly translates to improvements in the condition numbers of the GVM. Figure 4.2 depicts the growth of the condition numbers of the GVM in which Fekete nodes obtain the lowest numbers among the examined precision sets. Thus, by improving the condition numbers of the determinant of the GVM, Fekete nodes set a bound on the condition numbers of A and B.

4.3 $H(\nabla\wedge)$ Basis

With the $H(\nabla\wedge)$ basis, the situation is more complicated. The construction of the basis is by no means pure Lagrangian and developing a general bound analysis similar to that of section 4.2.1



(a) Order 7 (b) Order 8



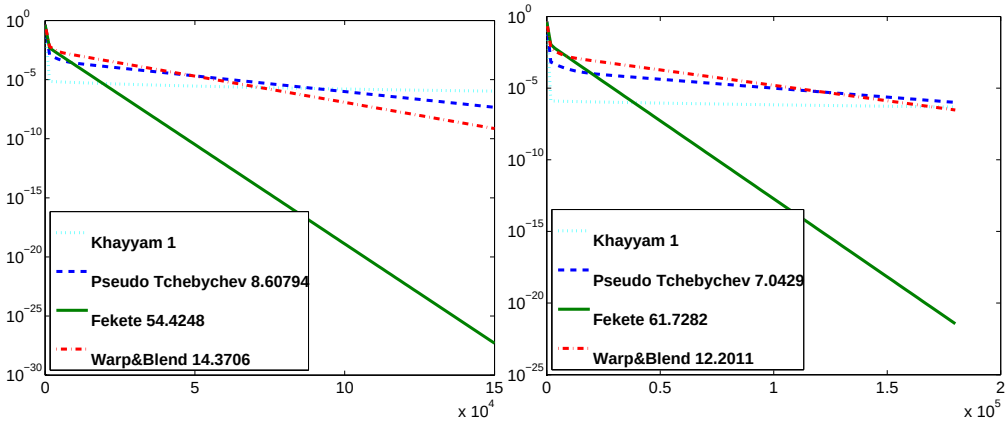
(c) Order 9 (d) Order 10

Figure 4.7: Figure 4.6 continued.

is a bit more difficult. Hence, a more empirical approach is used for the analysis of the effect of the choice of precision set on spectral FE matrices of problems posed int the $H(\nabla\wedge)$ space. Equations (4.21) and (4.22) provide the element-wise formulation typical of stiffness and mass matrices encountered in EM problems. In these equations, $\vec{\alpha}_i$ represent the components of the vector polynomial basis spanning the $H(\nabla\wedge)$ space [70, 86].

$$A = [A_{ij}] = \left[\int_{\Omega=\triangle} \nabla\wedge\vec{\alpha}_i.\nabla\wedge\vec{\alpha}_j \; d\Omega\right] \tag{4.21}$$

$$B = [B_{ij}] = \left[\int_{\Omega=\triangle} \vec{\alpha}_i.\vec{\alpha}_j \; d\Omega\right] \tag{4.22}$$



(a) Order 11 (b) Order 12

Figure 4.8: Figure 4.7 continued.

4.3.1 Basic Bound Analysis

For the sake of simplicity, we will consider the development of the required polynomial shape functions over the reference right triangle $\triangle = \{(x, y)|0 \leq x, y, 0 \leq x + y \leq 1\}$. Following the approach of [1, 12], the desired $H(\nabla\wedge)$ polynomial space can be spanned by the union the three bases defined in (4.23), (4.24) and (4.25) where Ω_{mn} are defined according to (4.26) , (4.27) and (4.28) and k_{ij} are the Koornwinder basis elements defined in (4.7).

$$\{k_{ij}(x, y)\Omega_{12}|_{x+y=1}\Omega_{12}|i + j = p, 0 \leq i, j\} \tag{4.23}$$

$$\{k_{ij}(x, y)\Omega_{13}|0 \leq i + j \leq p, 0 \leq i, j\} \tag{4.24}$$

$$\{k_{ij}(x, y)\Omega_{23}|0 \leq i + j \leq p, 0 \leq i, j\} \tag{4.25}$$

$$\Omega_{12} = x\hat{y} - y\hat{x} \tag{4.26}$$

$$\Omega_{13} = (y - 1)\hat{x} - x\hat{y} \tag{4.27}$$

$$\Omega_{23} = -y(-\hat{x}) + (x - 1)\hat{y} \tag{4.28}$$

Here we will denote the $H(\nabla\wedge)$ vector polynomial space defined on $\triangle$ by $\mathbf{P}^p(\triangle)$ where p stands for the order of the polynomial space. The dimension dm of the polynomial space $\mathbf{P}^p(\triangle)$ is given as:

$$dm \triangleq \dim(\mathbf{P}^p(\triangle)) = (p+1)(p+3) \tag{4.29}$$

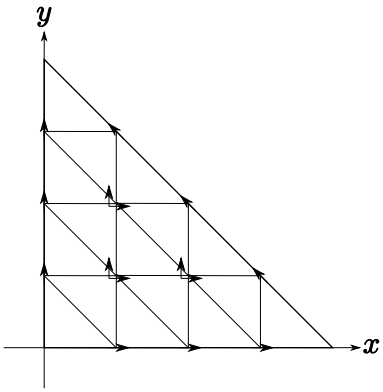


Figure 4.9: Graphical Representation of vector DoF (location and direction) for the 2^{nd} order-complete case.

Similar to the H^1 case, one starts by constructing a basis for the vector space dual to $\mathbf{P}^p(\triangle)$. In FEM literature, the components if this dual-space basis are regarded as Degree(s) of Freedom (DoF) [14]. In the H^1 case, the dual basis (or the Degree(s) of Freedom (DoF)) consist of simple polynomial evaluations at the interpolation points denoted by $|x_j$. The difference here, is that we are dealing with spatially 2-dimensional vector polynomials and that the desired basis for the dual space must comprise of polynomial evaluations [1] followed by an inner product with appropriate direction vectors, i.e. $|_{x_j}.\vec{e}_j$. In this way, the definition of GVM can be further extended to vector polynomial spaces. Figure 4.9, depicts a composition of the vector Degree(s) of Freedom (DoF) for the second order case expressed over KPI nodes. The general recipe is to have $(p+1)$ tangent Degree(s) of Freedom (DoF) on each edge and $(p)(p+1)$ Degree(s) of Freedom (DoF) (one along $\hat{x}$ and one along $\hat{y}$ at each internal node) inside $\triangle$. Hence, for the $H(\nabla\wedge)$ basis $\{\vec{\alpha}_i\}$, the GVM can be developed as (4.30) where $\vec{k}_j$ are the the $(p+1)(p+3)$ basis components defined in

[1] In general any linear map from the underlying scalar polynomial space to $\mathbb{R}$

(4.23), (4.24) and (4.25) linearly indexed with respect to j.

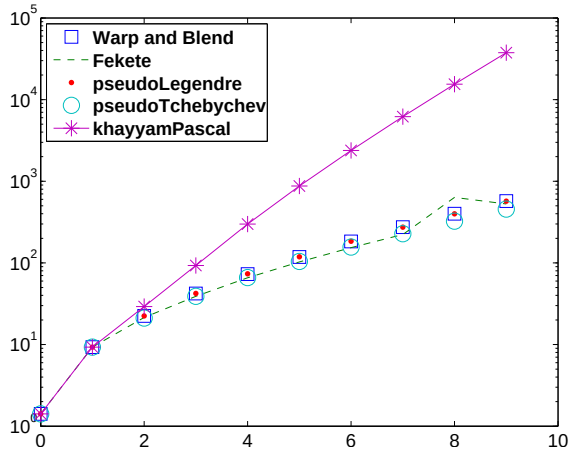


Figure 4.10: Condition number of the GVM ($H(\nabla\wedge)$) defined in (4.30) versus polynomial order on the right triangle $\triangle = \{(x,y)|0 \leq x,y, 0 \leq x+y \leq 1\}$.

$$V = [V_{ij}] \triangleq [\vec{k}_i(x_j).\vec{e}_j] \tag{4.30}$$

Using (4.30), the vector shape functions can now be derived as (4.31).

$$\vec{\alpha}_i(x) = \sum_{i=1}^{dm} L_{ij}\vec{k}_j(x) \therefore$$

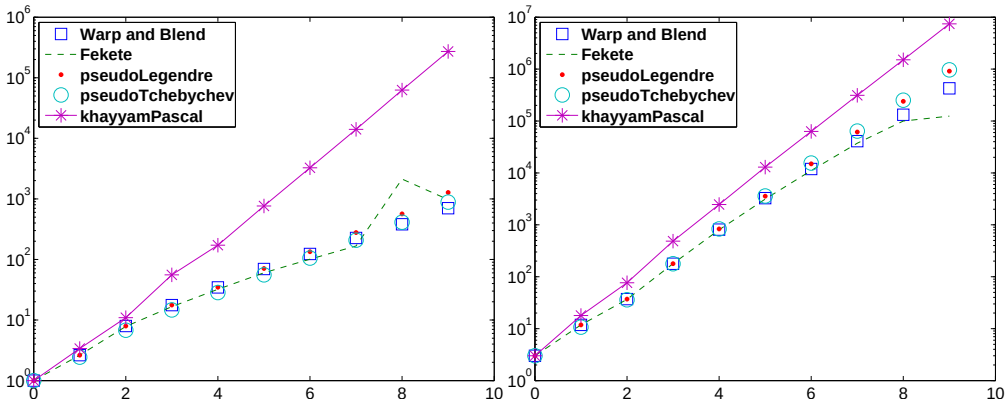
$$L = [L_{ij}] = V^{-1} \rightarrow \begin{bmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \\ \vdots \\ \vec{\alpha}_{dm} \end{bmatrix} = V^{-1} \begin{bmatrix} \vec{k}_1 \\ \vec{k}_2 \\ \vdots \\ \vec{k}_{dm} \end{bmatrix} \tag{4.31}$$

Using the derivation in (4.31) and implementing an approach similar to that of (4.10) and (4.11), A and B can be formulated as it follows in (4.32) and (4.33).

$$\begin{aligned} A = [A_{ij}] &= V^{-1}\left[\int_{\Omega=\triangle} \nabla\wedge\vec{k}_i.\nabla\wedge\vec{k}_j \, d\Omega\right] V^{-1T} = V^{-1}C_A V^{-1T} \\ \text{where } C_A &= \left[\int_{\Omega=\triangle} \nabla\wedge\vec{k}_i.\nabla\wedge\vec{k}_j \, d\Omega\right] \end{aligned} \tag{4.32}$$

$$B = [B_{ij}] = V^{-1}\left[\int_{\Omega=\triangle}\vec{k}_i.\vec{k}_j\ d\Omega\right]V^{-1T} = V^{-1}C_BV^{-1T}$$
$$\text{where } C_B = \left[\int_{\Omega=\triangle}\vec{k}_i.\vec{k}_j\ d\Omega\right] \tag{4.33}$$

From this point onwards, the analysis is very much similar to that of section 4.2.1. Hence, (4.19) holds for A and B matrices of (4.21) and (4.22). Thus, it is again concluded (similar to section 4.2.1) that bounding either the Frobenius or the l^2 induced condition numbers of the GVM bounds both the Frobenius and the l^2 induced condition numbers of A and B matrices defined in (4.21) and (4.22). Similar to the H^1 case, B is a full rank matrix while A and hence C_A are under determined [1, 58]. Note that the definition of C_A and C_B is independent from the choice of the precision set, i.e. the only precision set dependent part in (4.32) and (4.33) is the GVM, V. Figure 4.10 plots the condition number of the GVM (as a function of polynomial order) for various precision sets. Unlike in the H^1 case, Fekete nodes do not produce the lowest condition numbers among the examined precision sets. Similar to the H^1 case, however, KPI nodes lead to the highest GVM condition numbers among the examined precision sets.

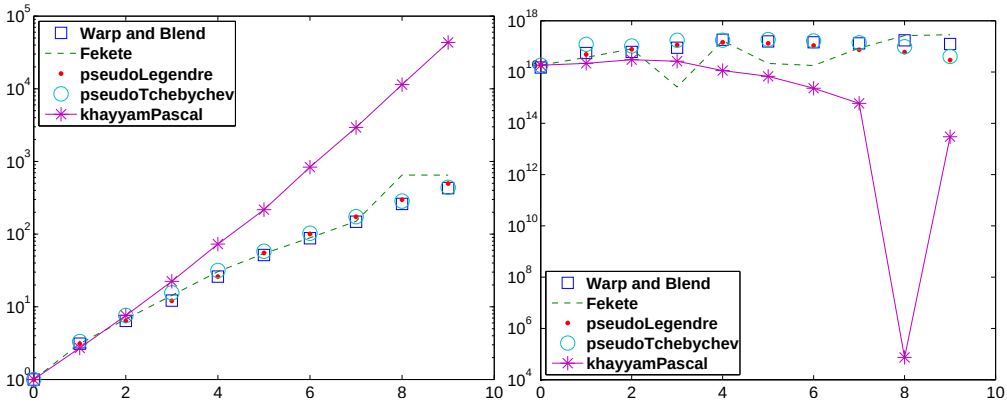


(a) stiffness matrix (b) mass matrix

Figure 4.11: Condition number versus polynomial order for the stiffness and mass matrices respectively defined in (4.21) and (4.22) on the right triangle $\triangle = \{(x, y)|0 \leq x, y, 0 \leq x+y \leq 1\}$

4.3.2 The Single Element Case

Similar to section 4.2.2, various precision sets are examined for their effect on the condition number of A and B matrices. The contrasting point, however, is that the quality of the so called



(a) Condition number (b) Separation Capability

Figure 4.12: Condition number and separation capability versus polynomial order for the element constraints matrix [1].

constraints matrix needs to be taken into consideration at the same time. The studied precision sets are the five sets mentioned earlier in section 4.1.2. Figure 4.11, depicts the condition number of single element stiffness and mass matrices for polynomial orders up to 9. From the figure, it is apparent that Fekete nodes generally provide better condition numbers for the mass matrix. Yet, it also indicates that in $H(\nabla\wedge)$, Fekete nodes no longer produce the lowest condition numbers for the stiffness matrix. Rather, there is a close competition between the examined nodal sets (except for KPI) in which warp & blend nodes dominate others at higher polynomial orders. Thus, from the overall performance of stiffness and mass matrices it can be concluded that warp & blend provide better condition numbers. This observation, however, seems to be incomplete since the effect of the choice of the precision set on the so called constraints matrix must also be taken into consideration. For this purpose, we shall first have a look at the condition numbers of the so called element constraints matrix. As defined in [1], the element constraints matrix C is a non rectangular matrices (for a p-order complete $H(\nabla\wedge)$ basis, the dimensions are $(p+1)(p+4)/2$ by $(p+1)(p+2)/2$). Thus the general definition of condition number defined in (4.13) must be applied. If the $H(\nabla\wedge)$ FEM matrices of (4.21) and (4.22) are applied for construction of an *eigenproblem* as formulated in (4.34), the proper null-space of A contributes to the introduction of a set of zero *eigenvalue* d solutions to (4.34).

$$A\,x = \lambda B\,x \tag{4.34}$$

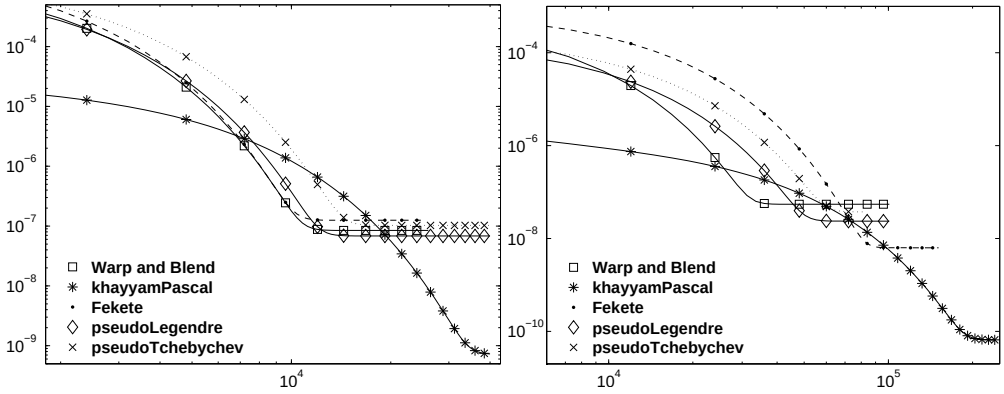
As discussed in [1, 56, 58], constraint equations can/must be used for elimination of such spurious modes. When dealing with HO spectral $H(\nabla\wedge)$ elements, this boils down to the incorporation of a linear *projection* operator $\mathcal{O}$ which basically projects any RHS vector down into the orthogonal complement of the null-space of A [1]. The orthogonal complement of $\mathcal{O}$ can be defined as $\mathcal{C} = I - \mathcal{O}^T$ where I is the appropriate identity matrix. Denoting the null-space of A by $\ker(A)$, it can be shown that the following orthogonality relations hold for any eigenvalue ν solving (4.34).

$$\begin{cases} \mathcal{O}^T\,B\,\nu = 0 : \nu \in \ker(A) \\ \mathcal{C}^T\,B\,\nu = 0 : \nu \notin \ker(A) \end{cases} \tag{4.35}$$

Thus, using (4.35) and considering the finite precision of machine floating point arithmetics, for every $\nu \in \ker(A)$, $\frac{\mathcal{C}^T B\,\nu}{\mathcal{O}^T B\,\nu}$ must be a significantly large number (it must be ∞ in the ideal case). In a similar manner, for every $\nu \notin \ker(A)$, $\frac{\mathcal{O}^T B\,\nu}{\mathcal{C}^T B\,\nu}$ must be a significantly large. Hence, we will take $\frac{\mathcal{C}^T B\,\nu}{\mathcal{O}^T B\,\nu}$ (for zero eigenvalued ν) or $\frac{\mathcal{O}^T B\,\nu}{\mathcal{C}^T B\,\nu}$ (for nonzero eigenvalued ν) as figures of the *separation* capability of the constraints matrix C from which $\mathcal{O}$ and $\mathcal{C}$ are assembled [1]. These values, can then be assorted in a vector which will be called the 'separation' capability vector. Figure 4.12b plots the *root mean square (RMS)* value of the separation vector for various precision sets. From figure 4.12b it is understood that the gap between separation capability of KPI nodes and other examined families of nodal set tends to grow with the polynomial order. Particularly at higher orders, warp & blend seem to yield stronger separation.

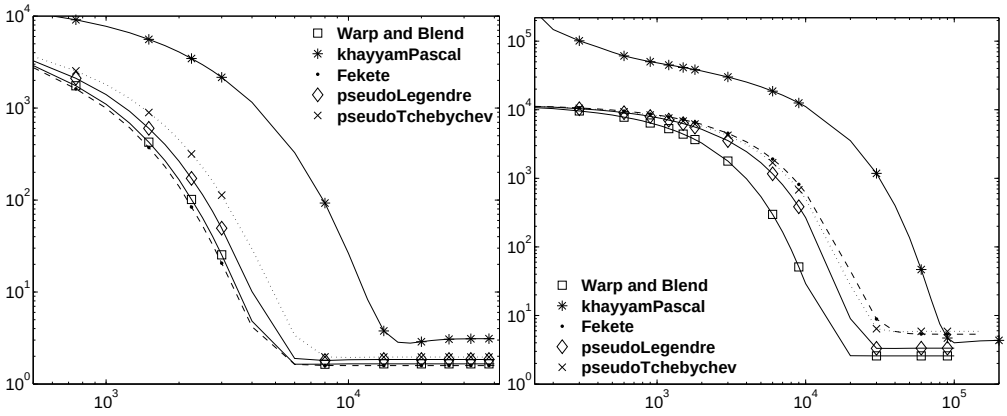
4.3.3 The General Case

Magnetostatics problems are the extreme case of EM problems where the angular frequency ω approaches zero. In such cases as magnetostatics problems and EM problems with small values of



(a) Order 7 elements (b) Order 9 elements

Figure 4.13: Krylov iteration history (solver error). All curves begin from a normalized error level equal to unity.



(a) Order 7 elements (b) Order 9 elements

Figure 4.14: l^2 error of the actual solution (as a function of the number of Krylov iterations) with respect to the analytical solution.

ω, the application of the constraints matrix to the solution of the final matrix problem is inevitable. Hence, to emphasize on the importance of the quality of the constraints matrix C, we will engage with the solution of a magnetostatics problem.

4.3.3.1 Problem Definition

The magnetostatic example presented here considers the 2D magnetic field solution of an infinitely long rectangular sectioned solenoid. The FEM process involves the solution of the weak form of the magnetic vector potential $\vec{A}$ such that $\vec{B}$ (or the magnetic flux density) equals

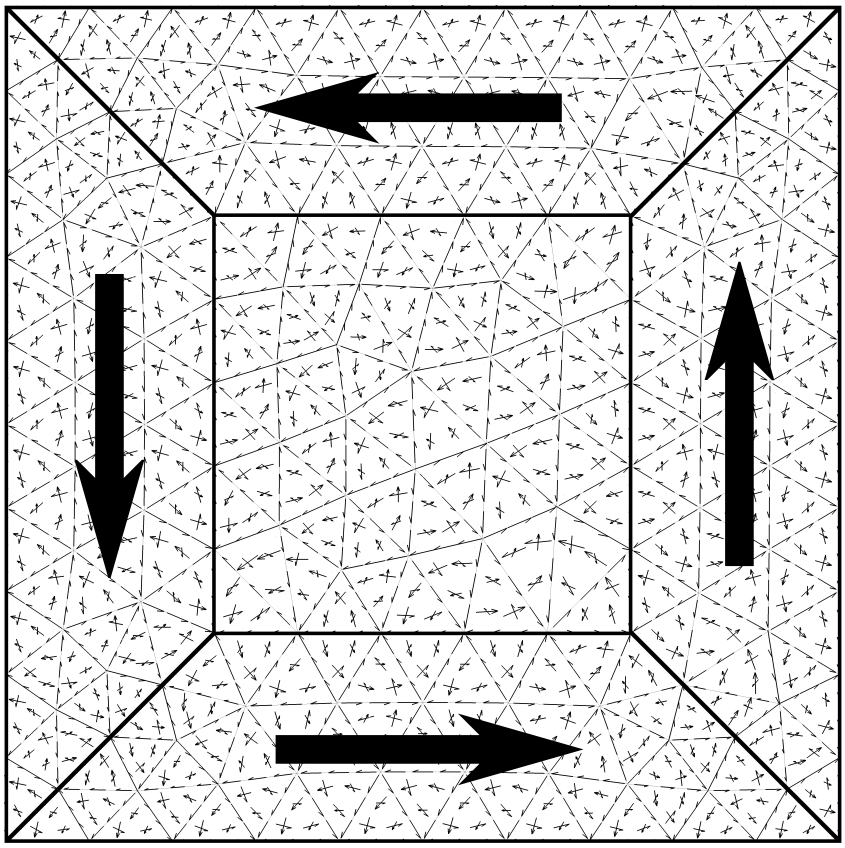


Figure 4.15: The 2^{nd}-order complete FE mesh with the associated $H(\nabla\wedge)$ DoF over the cross section of an infinitely long solenoid. The big arrows indicate the presence of electric current in regions. This example uses the Fekete nodes.

$\nabla \wedge \vec{\mathbf{A}}$. The associated FE mesh and geometry of the problem are depicted in figure 4.15. Denoted by big arrow signs in figure 4.15, the current carrying region of the solenoid is assumed to carry a constant current density J while a uniform magnetic permeability μ_0 is assumed over the entire problem domain. At the same time, the exterior boundary of the problem domain is terminated on a *perfect magnetic conductor (PMC)* which imposes the $\hat{\mathbf{n}} \wedge \frac{1}{\mu}\nabla \wedge \vec{\mathbf{A}} = 0$ condition. Equation (4.36) presents the final matrix form of the problem in which A and B matrices follow the definitions from (4.21) and (4.22). In (4.36), J denotes the vector that is produced by interpolation of the current density over the $H(\nabla\wedge)$ FE basis.

$$A\,\mathsf{x} = \mathsf{y}, \quad \mathsf{y} = B\,\mathsf{J} \tag{4.36}$$

4.3.3.2 Solution and Comparison

Using symmetry arguments and applying the Ampere's law it can be shown that the magnetic flux in the solenoid linearly grows from zero on the boundaries and reaches its maximum values on the edges of the central rectangular section; the magnetic flux intensity is constant across the central rectangular section.

Unlike the H^1 case, the quality of the interpolation precision sets cannot be evaluated by simple inspections of the Krylov solver error history. Due to the presence of a proper null-space of A, the RHS of the matrix equation must be projected into the range-space of the LHS matrix before a solution for the problem can sought [1]. In simple words, the part of the numerical deviation of the projected RHS that falls out of the range-space of the LHS matrix attributes to a solution error which cannot be resolved regardless of the number of Krylov iterations applied. Thus, the numerical accuracy of $\mathcal{O}$ sets a limit on how close the Krylov iterations can get to the true solution of the matrix problem. This is clearly observed in the plots of figure 4.13 where Krylov iterations reach a saturation at levels much higher than that of the machine precision saturation. In figure 4.13, it is observed that precision sets other than the KPI nodes possess higher convergence rates. Yet, except for convergence rates, relative error (the error norm prior to the first Krylov iteration is normalized to one) and the level of it's saturation are not indicative of the quality of the actual solution. Instead, as presented in figure 4.14, the l^2 error of the actual solution (with respect to the analytical solution) can be taken for a more meaningful inspection of the examined precision sets. In figure 4.14, it is observed that significant improvements in the solution accuracy and time can be achieved using precision sets other than the KPI nodes. For example, from figure 4.14b, it is understood that with warp & blend nodes, the same level of l^2 error can be achieved at 10 times less the number of iterations required for KPI nodes. Furthermore, the final (saturated) error levels can be considerably lower (more than twice in this case) with the choice of warp & blend nodes.

4.3.4 $H(\nabla\wedge)$ Conclusion

The observations made through this section indicate that the choice of the precision set has significant impact on the quality of FE matrices for $H(\nabla\wedge)$ problems. It is revealed that three factors must be taken into consideration for gauging the quality of a precision set with respect to FE problems in the $H(\nabla\wedge)$ space:

1. Condition number of the LHS matrix A.
2. Condition number of the RHS matrix B.
3. Condition number and accuracy of the constraints matrix C.

It must be emphasized that the reported observations are limited to polynomial orders less than 10, though warp & blend nodes lead to significant improvements in the performance of the the examined nodal sets. Hence, it is speculated that dedicated precision sets with better performance for $H(\nabla\wedge)$ problems can be constructed. The generalized definition of the GVM presented in section 4.3.1 suggests that (in analogy to [24]) improved nodal sets can be constructed based on maximization of this matrix. For example, [24] uses a time evolution of the configuration of the nodal set in which the nodes are iteratively moved toward the maxima of the Lagrangian functions associated with them. The approach of [24] is currently only applied to the H^1 spectral basis over the triangle. In the $H(\nabla\wedge)$ case, however, two Degree(s) of Freedom (DoF) are associated to each internal node and thus a direct generalization of [24] would not be possible. Rather, it is speculated that the pair of Degree(s) of Freedom (DoF) currently defined at each internal node of $\triangle$ can be associated to a pair of potentially distinct points. This may form the basis for future developments of improved precision sets dedicatedly designed for the $H(\nabla\wedge)$ basis.

Chapter 5

Continuous Material Property Elements

5.1 Motivation

The advent of technologies such as 3D photo-lithography and holography has lead to accurate fabrication of devices such as the Luneburg lens. Furthermore, many real life applications involve spatial changes in *material property tensors (MPTs)* . Clearly, integration of CIMPTs inside individual FEs adds to the flexibility and efficiency of FEMs. Curved FEs have been extensively used to mitigate geometrical non-conformities associated to rectilinear approximation of curvilinear features while CIMPT are conventionally handled by element-wise constant approximation. FE matrices are traditionally evaluated through 1) numerical cubature or 2) *universal matrices (UMs)* In essence, both methods rely on polynomial integration. Furthermore, complications associated with evaluation of FE matrices on elements with curved boundaries or CIMPTs are identical in nature, i.e. integrals with non-constant Jacobian or MPT terms. Alternatively, in this work, a generalized UM approach is proposed, which reduces the cost associated with evaluation of FE matrices with curvilinear features and CIMPTs. Motivated by simulation of a non-graded *Luneburg* lens, the conventional element-wise constant MPTs are replaced with polynomial rep-

resentations yielding the following benefits: **a)** more accurate physical models for problems with CIMPTs and curved features; and, **b)** better utilization of computer resources when higher-order curved elements replace lower-order elements of smaller physical dimensions.

5.2 Introduction

Curved FEs have for long been used to mitigate geometrical-model non-conformities encountered in computer aided modeling and simulation of real life problems. Iso-parametric FEs provide a flexible tool for discretization of curved geometries by means of polynomial or rational approximation of element coordinate-transform dependent terms [87, 88]. In general, both numerical cubature and UM methods have been deployed for evaluation of FE matrix entries. Nevertheless, both numerical cubature and UMs are theoretically based on polynomial approximation of the integrand, which suggests that they should be equally applicable to various kinds of FE matrix evaluations. Moreover, in conventional FEM codes, material properties, e.g. permittivity and permeability tensors in EM, are treated as element-wise constant functions while some recent works have examined the idea of having non-constant material properties within individual elements [2] verifying that the techniques can be advantageous. Also in [49], Webb uses a polynomial representation of the magnetic MPT for evaluation of FE matrices arising from a nonlinear magnetic problem. While element-wise constant material properties can still be used in a wide variety of problems, there are cases where material properties do change continuously in space. The Luneburg lens, Maxwell's fisheye lens, semiconductor devices and media where material properties are affected or governed by diffusion of heat, moisture, dopant etc. are practical examples of problems where the conventional element-wise constant material approach may fail to provide an accurate model of the underlying physics. In cases as such, material properties follow continuous spatial fluctuations that can be modeled more accurately using polynomial approximation, e.g. interpolation.

In this work, we present a generalized UM, i.e. *umiversal array (UA)* , approach for evalu-

ation of FE matrices for elements in which material properties are allowed to vary continuously across individual elements (section 5.3) and where element boundaries are allowed to conform to given curvature. The method can be equally applicable to conventional rectilinear elements with constant material properties, in which it simplifies to the well known UM approach [48] with the caveat that our symmetric evaluation of matrix entries further reduces the computational complexity.

As a test to the presented approach, we will implement a conformal FEs domain decomposition method (FEM-DDM) [89] solution of an ungraded Luneburg lens operated in the X-Band. The actual DDM formulation is briefly discussed in section 5.4. The lens problem is an exterior radiation/propagation problem with continuously changing material properties and curved boundaries associated with the surface of the spherical Luneburg lens.

Accurate expression for conformal *perfectly matched layer (PML)* 's [90] permittivity and permeability tensors involves continuous functions of the spatial coordinates [91]. Hence, a PML with continuously varying material proprieties is examined for the exterior truncation of the wave problem [92]. Obviously, the resulting FE matrices are, again, evaluated using the proposed CIMPT method. The PML formulation is briefly discussed in section 5.5. Finally, numerical results are presented in section 5.6.

5.2.1 Methodology

Systematically, the evaluation process of FE stiffness and mass matrices can be cast into the following basic steps:

1. Assume a reference element with a fixed geometry, e.g. figure 5.1a.

2. Develop the required polynomial/vector-polynomial basis on the reference element.

3. Develop the required transformation rules for the basis, its derivations and the Jacobian defined between the reference element and a presumed physical element, e.g. figure 5.1b.

4. Express the required integral-differential form on the physical element.

5. Pull back the integration operation onto the reference element, i.e. re-express the integral over the reference element.

6. Identify the factors that are solely and entirely determined by the geometry of the physical element from those solely defined (depend) on the reference element. Hereafter, we shall refer to the identified physical-element geometry dependent factors as the metric factors of the integrand.

7. Being independent from the reference element coordinates, the metric factors are pulled out of the integral (over the reference element) while the remaining parts become independent of the metric properties of the physical element. Hence, the integrals can be pre-calculated and stored in the so called UAs and the evaluation of FE matrices turns into a sequence of multiply-add operations.

The following section explains how the abovementioned steps are adopted for evaluation of FE matrices associated with time harmonic Maxwell's equations. The FE matrices are then used as an integral part of a conformal-DDM code which, for the purpose of verification, is used for the solution of a few example problems including an ungraded Luneburg lens.

5.2.2 Notation

Throughout this work, whenever summation bounds of are not explicitly given, it is implied that they fall in the $0 \leq index < 3$ bounds. Furthermore the superscript $^{\mathsf{T}}$ is used to denote matrix transposition. At the same time, column vectors and matrices are denoted by $[a_i]$ and $[a_{ij}]$ while their individual entries are denoted as $[a_{ij}]_{ij}$ and $[a_i]_i$, respectively. The J and K symbols are used to denote forward and backward Jacobian matrices between the reference and the actual physical elements; and, F denotes a coordinate mapping defined between them. As visualized in figure 5.1, the reference and the physical elements are respectively denoted by $\mathcal{K}_r$ and $\mathcal{K}_p$

while ξ and x superscripts stand for the coordinate systems defined on $\mathcal{K}_r$ and $\mathcal{K}_p$. The $\times$ sign represents matrix multiplication while, in order to avoid confusion with $\times$, $\wedge$ is used to denote exterior products in $\mathbb{R}^3$. Throughout the article, $\nabla^{\boldsymbol{x}}$ and ∇^{ξ} denote the gradient operators in their respective coordinates, i.e. x and ξ. Hence, for any sufficiently smooth real function ϱ defined on either the reference or the physical element the column vector form of the gradient is denoted as $[\nabla]^{\boldsymbol{x}}\varrho \triangleq [\frac{\partial \varrho}{\partial x_i}]$ or $[\nabla]^{\boldsymbol{\xi}}\varrho \triangleq [\frac{\partial \varrho}{\partial \xi_i}]$. Furthermore, various physical quantities and functional spaces are used in this work(particularly in section 5.4) for which a brief listing can be found in Table 5.1.

5.3 Universal Arrays

This section presents a step-by-step realization of the concepts introduced in section 5.2.1.

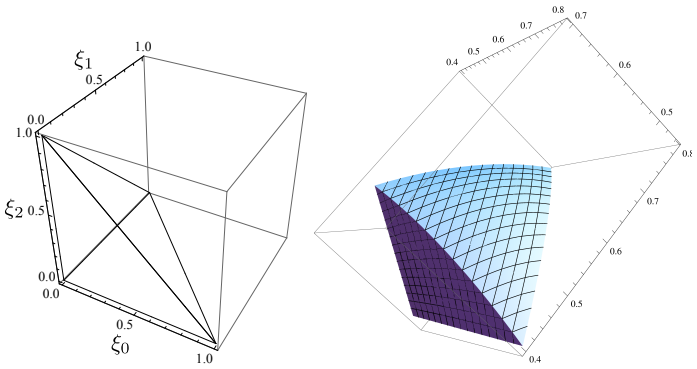
5.3.1 The Reference Element $\mathcal{K}_r$

As depicted in figure 5.1a, the reference tetrahedron is chosen as $\mathcal{K}_r = \{(\xi_0, \xi_1, \xi_2, \xi_3) | \xi_0 + \xi_1 + \xi_2 + \xi_3 = 1, 0 \leq \xi_0, \xi_1, \xi_2, \xi_3 \leq 1\}$. Note that the four coordinate variables on $\mathcal{K}_r$ are dependent through $\xi_0 + \xi_1 + \xi_2 + \xi_3 = 1$. Hence, ξ_3 is often omitted and replaced by $1 - (\xi_0 + \xi_1 + \xi_2)$. Monomial and polynomial integrations on $\mathcal{K}_r$ can be handled analytically.

Given the coordinate mapping $\digamma : (\xi_0, \xi_1, \xi_2) \rightarrow (x_0, x_1, x_2)$ between $\mathcal{K}_r$ and a physical element of interest $\mathcal{K}_p$, the Jacobian matrices J and K can be defined as in (5.1) in which $\circ$ denotes function composition. It is thus obvious that J and K are the inverse of each other and that $\det \mathsf{J} \neq 0$ and $\det \mathsf{K} \neq 0$ are guaranteed by definition.

$$\mathsf{J} = \left[\frac{\partial x_i}{\partial \xi_j}\right] = \left[\frac{\partial\, x_i \circ \digamma(\boldsymbol{\xi})}{\partial \xi_j}\right] \leftrightarrow \mathsf{K} = \left[\frac{\partial \xi_i}{\partial x_j}\right] = \left[\frac{\partial\, \xi_i \circ \digamma^{-1}(\boldsymbol{x})}{\partial x_j}\right] \tag{5.1}$$

Straightforward application of the chain rule reveals that the following relations hold between column vector representation of the gradients defined over $\mathcal{K}_r$ and $\mathcal{K}_p$.



(a) The 3D reference element. (b) An actual curved element.

Figure 5.1: A visualization of the concept of reference/physical elements.

$$[\nabla]^{\boldsymbol{x}}\varrho = \mathsf{K}^{\mathsf{T}} \times [\nabla^{\boldsymbol{\xi}}]\varrho \tag{5.2}$$

$$[\nabla]^{\boldsymbol{\xi}}\varrho = \mathsf{J}^{\mathsf{T}} \times [\nabla^{\boldsymbol{x}}]\varrho \tag{5.3}$$

5.3.2 The $H(curl, \mathcal{K}_r)$ Basis and Transformations

Construction of the $H(curl, \mathcal{K}_r)$ conforming basis (in either of its hierarchical or spectral forms) has been extensively discussed in the literature [1, 12, 56, 65, 66]. Regardless of the choice of the actual basis, any (vector) basis function $\vec{\beta}_u^{\boldsymbol{\xi}}$ (over the reference element) can be expressed as the sum of three scalar polynomials $\{\beta_u^{\boldsymbol{\xi}} | 0 \leq i < 3\}$ each multiplied by a direction factor $\{\nabla^{\boldsymbol{\xi}}\xi_i | 0 \leq i < 3\}$ [56]. Expressed explicitly, we have

$$\vec{\beta}_u^{\boldsymbol{\xi}} = \sum_j \beta_{u,j}^{\boldsymbol{\xi}} \nabla^{\boldsymbol{\xi}}\xi_j \tag{5.4}$$

where the u subscript signifies a particular basis function.

From (5.4), it is obvious that a basis function $\vec{\beta}_u^{\boldsymbol{\xi}}$ on $\mathcal{K}_r$ can be represented by a column vector representation $[\beta_{u,i}^{\boldsymbol{\xi}}]$ with respect to the $\{\nabla^{\boldsymbol{\xi}}\xi_i\}$ direction basis. It is not difficult to observe that replacing the $\nabla^{\boldsymbol{\xi}}$ term in (5.4) with $\nabla^{\boldsymbol{x}}$ yields the $H(curl, \mathcal{K}_p)$ conforming basis on the physical

element $\mathcal{K}_p$. In other words, since $\boldsymbol{\xi}$ can be written as a function of $\boldsymbol{x}$ (and vice versa), basis functions derived as (5.5) satisfy the necessary tangential continuity conditions for a $H(curl, \mathcal{K}_r)$ conforming basis on $\mathcal{K}_p$ [93].

$$\vec{\beta}_u^{\boldsymbol{x}} = \sum_j \beta_{u,j}^{\boldsymbol{\xi}} \nabla^{\boldsymbol{x}} \xi_j \tag{5.5}$$

From (5.5), one writes:

$$[\beta_{u,i}^{\boldsymbol{x}}]_i = \sum_j \left[\frac{\partial \xi_j}{\partial x_i}\right]_i \beta_{u,j}^{\boldsymbol{\xi}} \rightarrow \left[\beta_{u,i}^{\boldsymbol{x}}\right] = \mathsf{K}^{\mathsf{T}} \times \left[\beta_{u,j}^{\boldsymbol{\xi}}\right] \tag{5.6}$$

Hence, with respect to the $\{\nabla^{\boldsymbol{x}} x_i\}$ direction basis, the column vector representation of $\vec{\beta}_u^{\boldsymbol{x}}$, i.e. $\left[\beta_{u,i}^{\boldsymbol{x}}\right]$, is obtained as $\mathsf{K}^{\mathsf{T}} \times [\beta_{u,j}^{\boldsymbol{\xi}}]$.

5.3.3 The Mass Matrix

Using the electrical field formulation, the entries of the FE mass matrix T for an electromagnetic radiation/propagation problem can be formulated as

$$\mathsf{T}_{uv} = \int_{\mathcal{K}_p} \vec{\beta}_u^{\boldsymbol{x}} \cdot \ (\bar{\bar{\varepsilon}} \cdot \vec{\beta}_v^{\boldsymbol{x}}) d\mathcal{K}_p = \int_{\mathcal{K}_p} [\beta_{u,i}^{\boldsymbol{x}}]^{\mathsf{T}} \times \bar{\bar{\varepsilon}} \times [\beta_{v,j}^{\boldsymbol{x}}] d\mathcal{K}_p \tag{5.7}$$

where $[\beta_{u,i}^{\boldsymbol{x}}]$ and $\bar{\bar{\varepsilon}}$ respectively denote the column vector representation of $\vec{\beta}_u^{\boldsymbol{x}}$ (discussed in section 5.3.2) and the permittivity tensor.

Using (5.6) and pulling the integration back onto $\mathcal{K}_r$, (5.7) can be written as:

$$\mathsf{T}_{uv} = \int_{\mathcal{K}_r} [\beta_{u,i}^{\boldsymbol{\xi}}]^{\mathsf{T}} \times \mathsf{K} \times \bar{\bar{\varepsilon}} \times \mathsf{K}^{\mathsf{T}} [\beta_{v,j}^{\boldsymbol{\xi}}] \det \mathsf{J} \ d\mathcal{K}_r \tag{5.8}$$

Next, if we define $[\varsigma_{ij}]$ as (5.9), T is obtained as (5.10).

$$[\varsigma_{ij}] \triangleq \det \mathsf{J}\,\mathsf{K}\times\bar{\bar{\varepsilon}}\times\mathsf{K}^{\mathsf{T}} \tag{5.9}$$

$$\mathsf{T}_{uv} = \sum_{i,j} \int_{\mathcal{K}_r} [\beta^{\boldsymbol{\xi}}_{u,i}]_i \, [\varsigma_{ij}]_{ij} \, [\beta^{\boldsymbol{\xi}}_{v,j}]_j \, d\mathcal{K}_r \tag{5.10}$$

Now, if $[\varsigma_{ij}](\boldsymbol{x})$ is constant across an element, it can be pulled out of the integral. Otherwise, $[\varsigma_{ij}](\boldsymbol{x})$ can be interpolated using a Lagrangian polynomial basis $\{\ell_s(\boldsymbol{\xi})\}$. Namely,

$$[\varsigma_{ij}](\boldsymbol{x}) \approx \sum_s [\varsigma_{ij}]|_{\boldsymbol{x}=\boldsymbol{x}_s} \ell_s(\boldsymbol{\xi} \circ F^{-1}(\boldsymbol{x})) \tag{5.11}$$

In (5.11), $\{\ell_s\}$ denotes the appropriate Lagrangian basis functions corresponding to an interpolation nodal set $\{\boldsymbol{\xi}_s\}$. It is clear that the Lagrangian polynomials are originally defined over $\mathcal{K}_r$ and transformed onto $\mathcal{K}_p$. Hence, given the interpolation nodal set $\{\boldsymbol{\xi}_s\}$ on $\mathcal{K}_r$, the interpolation can be built from interpolant values at the corresponding physical node set $\{\boldsymbol{x}_s | \boldsymbol{x}_s = F(\boldsymbol{\xi}_s)\}$. Therefore, for the sake of compactness we define ς_{ijs} to represent as $[\varsigma_{ij}]|_{\boldsymbol{x}=\boldsymbol{x}_s}$ and have:

$$\varsigma_{ijs} \triangleq [\varsigma_{ij}]_{ij}\big|_{\boldsymbol{x}=\boldsymbol{x}_s} = \Big(\det \mathsf{J}\,\mathsf{K}\times\bar{\bar{\varepsilon}}\times\mathsf{K}^{\mathsf{T}}\Big)\Big|_{\boldsymbol{x}=\boldsymbol{x}_s} \tag{5.12}$$

Hence, defining $\varkappa_{ui,vj}$ as (5.13) we end up with (5.14) in which S stands for the number of interpolation points or $\dim\{\ell_s\}$ equivalently. In this work, a 2^{nd} and a 5^{th}-order equidistant interpolation nodal sets, i.e $S = 10$ and $S = 56$, have been used for rectilinear and curvilinear elements, respectively. However, it is clear that the order of the polynomial basis $\{\ell_s\}$ can be increased/decreased based upon necessity. In other words, UAs with various polynomial orders for the Lagrangian $\{\ell_s\}$ basis can be precalculated, stored and used based on the geometrical/material complexity of individual elements.

$$\varkappa_{ui,vj} \triangleq [\beta^{\boldsymbol{\xi}}_{u,i}]_i [\beta^{\boldsymbol{\xi}}_{v,j}]_j \tag{5.13}$$

$$\mathsf{T}_{uv} = \sum_{0\le s<S} \sum_{i,j} \varsigma_{ijs} \int_{\mathcal{K}_r} \varkappa_{ui,vj} \ell_s(\boldsymbol{\xi})\, d\mathcal{K}_r \tag{5.14}$$

Next, our attention is turned to the exploitation of the algebraic symmetries of (5.14). First it is observed that if the integral part of (5.14) is precalculated, evaluation of each T_{uv} requires $9\times 9S$ multiply-add operations which is due to 9 operations for each ς_{ijs} entry times 9 operations for the summation in the ij indices as formulated in (5.14). However, closer examination of (5.13) reveals that $\varkappa_{ui,vj} = \varkappa_{vj,ui}$. This suggests that evaluation of $(\mathsf{T}_{uv}+\mathsf{T}_{vu})/2$ and $(\mathsf{T}_{uv}-\mathsf{T}_{vu})/2$ instead of T_{uv} and T_{vu} can be achieved with reduced computational complexity. Hence, using some algebraic manipulation the more symmetric formulation of (5.15) through (5.18) is obtained:

$$\mathsf{T}_{uv}\pm\mathsf{T}_{vu} = \sum_{i\le j,0\le s<S} (\varsigma_{ijs}\pm\varsigma_{jis}) \int_{\mathcal{K}_r} \frac{\varkappa_{ui,vj}\pm\varkappa_{uj,vi}}{1+\delta_{ij}} \ell_s d\mathcal{K}_r \tag{5.15}$$

$$\overset{\circ}{\mathsf{T}}{}^{\pm uivjs} \triangleq \int_{\mathcal{K}_r} \frac{\varkappa_{ui,vj}\pm\varkappa_{uj,vi}}{2+2\delta_{ij}} \ell_s\, d\mathcal{K}_r \tag{5.16}$$

$$\mathsf{C}^{\pm}_{ijs} \triangleq \varsigma_{ijs}\pm\varsigma_{jis} \tag{5.17}$$

$$\frac{1}{2}(\mathsf{T}_{uv}\pm\mathsf{T}_{vu}) = \sum_{i\le j,0\le s<S} \mathsf{C}^{\pm}_{ijs} \quad \overset{\circ}{\mathsf{T}}{}^{\pm uivjs} \tag{5.18}$$

Implementation of the above formulation leads to total $\frac{9\times(6+3)}{2}S$ multiply-add operations per matrix entry as it seeks to find two T entries. i.e. $\mathsf{T}_{uv}+\mathsf{T}_{vu}$ and $\mathsf{T}_{uv}-\mathsf{T}_{vu}$, at once and because C^{-}_{ijs} is identical to zero for cases where $i=j$.

5.3.4 The Stiffness Matrix

Similar to section 5.3.3, the electrical field formulation for electromagnetic radiation/propagation problem yields the following expression for the FE stiffness matrix S:

$$\mathsf{S}_{uv} = \int_{\mathcal{K}_p} \nabla^{\boldsymbol{x}} \wedge \vec{\beta}_u^{\boldsymbol{x}} \cdot \ (\bar{\bar{\mu}}^{-1} \nabla^{\boldsymbol{x}} \wedge \vec{\beta}_v^{\boldsymbol{x}}) d\mathcal{K}_p \tag{5.19}$$

In (5.19), $\bar{\bar{\mu}}$ represents the magnetic permeability tensor. Let's take $\vec{\beta}_v^{\boldsymbol{x}}$ as one of the vector polynomial basis over $\mathcal{K}_p$ and expand it using (5.5). We would like to express the curl of $\vec{\beta}_u^{\boldsymbol{x}}$ in a handy form. First, we take advantage of the vector identity $\nabla \wedge (f\vec{F}) = \nabla f \wedge \vec{F} + f\nabla \wedge \vec{F}$ and the fact that $\nabla \wedge (\nabla \phi) = 0$ for any sufficiently smooth function ϕ:

$$\begin{aligned} \nabla^{\boldsymbol{x}} \wedge \vec{\beta}_v^{\boldsymbol{x}} &= \sum_m \nabla^{\boldsymbol{x}} [\beta_{v,m}^{\boldsymbol{\xi}}]_m \wedge \nabla^{\boldsymbol{x}} \xi_m + [\beta_{v,m}^{\boldsymbol{\xi}}]_m \nabla^{\boldsymbol{x}} \wedge \nabla^{\boldsymbol{x}} \xi_m \\ &= \sum_m \nabla^{\boldsymbol{x}} [\beta_{v,m}^{\boldsymbol{\xi}}]_m \wedge \nabla^{\boldsymbol{x}} \xi_m \end{aligned} \tag{5.20}$$

Knowing that $\nabla^{\boldsymbol{x}} f = \sum_m \frac{\partial f}{\partial \xi_m} \nabla^{\boldsymbol{x}} \xi_m$ we get:

$$\nabla^{\boldsymbol{x}} \wedge \vec{\beta}_v^{\boldsymbol{x}} = \sum_{m,n} \frac{\partial [\beta_{v,n}^{\boldsymbol{\xi}}]_n}{\partial \xi_m} \nabla^{\boldsymbol{x}} \xi_m \wedge \nabla^{\boldsymbol{x}} \xi_n \tag{5.21}$$

Since $\nabla^{\boldsymbol{x}} \xi_m \wedge \nabla^{\boldsymbol{x}} \xi_n = -\nabla^{\boldsymbol{x}} \xi_n \wedge \nabla^{\boldsymbol{x}} \xi_m$, it further develops that:

$$\nabla^{\boldsymbol{x}} \wedge \vec{\beta}_v^{\boldsymbol{x}} = \sum_{m<n} \left(\frac{\partial [\beta_{v,n}^{\boldsymbol{\xi}}]_n}{\partial \xi_m} - \frac{\partial [\beta_{v,m}^{\boldsymbol{\xi}}]_m}{\partial \xi_n} \right) \nabla^{\boldsymbol{x}} \xi_m \wedge \nabla^{\boldsymbol{x}} \xi_n \tag{5.22}$$

In (5.22) the summation over m, n is replaced with $m < n$ which is due to the anti-symmetric nature of the exterior product in $\nabla^{\boldsymbol{x}} \xi_m \wedge \nabla^{\boldsymbol{x}} \xi_n$. Next, given the $\bar{\bar{\mu}}^{-1}$ tensor as $\bar{\bar{\mu}}^{-1} = \sum_{p,q} \hat{x}_p \bar{\bar{\mu}}_{p,q}^{-1} \hat{x}_q$ we expand $\bar{\bar{\mu}}^{-1} \cdot \nabla^{\boldsymbol{x}} \wedge \vec{\beta}_v^{\boldsymbol{x}}$ as:

$$\begin{aligned} &\bar{\bar{\mu}}^{-1} \cdot (\nabla^{\boldsymbol{x}} \wedge \vec{\beta}_v^{\boldsymbol{x}}) = \\ &\sum_{p,q} \hat{x}_p \bar{\bar{\mu}}_{pq}^{-1} \sum_{m<n} \left(\frac{\partial [\beta_{v,n}^{\boldsymbol{\xi}}]_n}{\partial \xi_m} - \frac{\partial [\beta_{v,n}^{\boldsymbol{\xi}}]_n}{\partial \xi_m} \right) (\hat{x}_q \cdot \nabla^{\boldsymbol{x}} \xi_m \wedge \nabla^{\boldsymbol{x}} \xi_n) \end{aligned} \tag{5.23}$$

Now following (5.19), in order to evaluate $\nabla^{\boldsymbol{x}} \wedge \vec{\beta}_u^{\boldsymbol{x}} \cdot \bar{\bar{\mu}}^{-1} \cdot \nabla^{\boldsymbol{x}} \wedge \vec{\beta}_v^{\boldsymbol{x}}$, $\varkappa_{uij,vmn}$ and $\varsigma_{ijp,mnq}$ are

defined in the following equations:

$$\varsigma_{ijp,mnq} \triangleq (\nabla^{\boldsymbol{x}}\xi_i \wedge \nabla^{\boldsymbol{x}}\xi_j \cdot \hat{x}_p)(\nabla^{\boldsymbol{x}}\xi_m \wedge \nabla^{\boldsymbol{x}}\xi_n \cdot \hat{x}_q) \tag{5.24}$$

$$\varkappa_{uij,vmn} \triangleq \left(\frac{\partial[\beta^{\boldsymbol{\xi}}_{u,j}]_j}{\partial\xi_i} - \frac{\partial[\beta^{\boldsymbol{\xi}}_{u,i}]_i}{\partial\xi_j}\right)\left(\frac{\partial[\beta^{\boldsymbol{\xi}}_{v,n}]_n}{\partial\xi_m} - \frac{\partial[\beta^{\boldsymbol{\xi}}_{v,m}]_m}{\partial\xi_n}\right) \tag{5.25}$$

Hence, it is important to observe the following symmetries for $\varkappa_{uij,vmn}$ and $\varsigma_{ijp,mnq}$:

$$\varkappa_{uij,vmn} = \varkappa_{vmn,uij} \tag{5.26}$$

$$\varsigma_{ijp,mnq} = \varsigma_{mnq,ijp} \tag{5.27}$$

$$\varsigma_{ijp,mnq}\varkappa_{uij,vmn} = \varsigma_{mnq,ijp}\varkappa_{vmn,uij} \tag{5.28}$$

In order to fully exploit the abovementioned symmetries, one must head for the evaluation of $\frac{1}{2}(\mathsf{S}_{uv} \pm \mathsf{S}_{vu})$ instead of S_{uv}.

$$\mathsf{S}_{uv} \pm \mathsf{S}_{vu} = \sum_{\substack{i<j,p \\ m<n,q}} \int_{\mathcal{K}_r} \bar{\bar{\mu}}^{-1}_{pq} \varsigma_{ijp,mnq}(\varkappa_{uij,vmn} \pm \varkappa_{vij,umn}) \det \mathsf{J}\, d\mathcal{K}_r \tag{5.29}$$

At this point, we turn our attention to the $\bar{\bar{\mu}}^{-1}_{pq}\varsigma_{ijp,mnq}\det \mathsf{J}$ part of (5.29). This term bares all the physical information on the geometry and material properties of the element. Hence, as a function of $\boldsymbol{x}$, it can be replaced with an appropriate Lagrangian interpolation similar to what was used in section section 5.3.3. Thus, as reflected in (5.33), an 's' subscript is added to $\varsigma_{ijp,mnq}$ and the summation indices. Namely, we write $\varsigma_{ijp,mnq,s}$ with the 's' subscript signifies evaluation at

$\boldsymbol{x} = \boldsymbol{x}_s$. In other words we have:

$$\bar{\bar{\mu}}^{-1}_{pqs} \triangleq \bar{\bar{\mu}}^{-1}_{pq}(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}_s} \tag{5.30}$$

$$\varsigma_{ijp,mnq,s} \triangleq \varsigma_{ijp,mnq}(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}_s} \tag{5.31}$$

$$\det \mathsf{J}_s \triangleq \det \mathsf{J}|_{\boldsymbol{x}=\boldsymbol{x}_s} \tag{5.32}$$

Hence, (5.29) is reformulated as:

$$\mathsf{S}_{uv} \pm \mathsf{S}_{vu} = \sum_{\substack{i<j,\,p \\ m<n,\,q \\ 0 \le s < S}} \int_{\mathcal{K}_r} \bar{\bar{\mu}}^{-1}_{pqs} \varsigma_{ijp,mnq,s} \det \mathsf{J}_s \ell_s(\boldsymbol{\xi})(\varkappa_{uij,vmn} \pm \varkappa_{vij,umn}) d\mathcal{K}_r \tag{5.33}$$

Now, taking advantage of the $(ij) \leftrightarrow (mn)$ symmetry, (5.33) is written as (5.34).

$$\mathsf{S}_{uv} \pm \mathsf{S}_{vu} = \sum_{\substack{i<j,\,p \\ m<n,\,q \\ ij \le mn \\ 0 \le s < S}} \int_{\mathcal{K}_r} \frac{\varkappa_{uij,vmn} \pm \varkappa_{vij,umn}}{1+\delta_{ij,mn}} \bar{\bar{\mu}}^{-1}_{pqs}(\varsigma_{ijp,mnq,s} \pm \varsigma_{mnp,ijq,s}) \ell_s(\boldsymbol{\xi}) \det \mathsf{J}_s d\mathcal{K}_r \tag{5.34}$$

Except for the $\bar{\bar{\mu}}^{-1}_{pqs}$ term, the above representation is symmetric in the $p \leftrightarrow q$ sense. Hence, the

symmetry can be further exploited to yield (5.35).

$$\mathsf{S}_{uv} \pm \mathsf{S}_{vu} = \sum_{\substack{i<j \\ m<n \\ ij \le mn \\ p \le q \\ 0 \le s < S}} \int_{\mathcal{K}_r} \frac{\varkappa_{uij,vmn} \pm \varkappa_{vij,umn}}{1+\delta_{ij,mn}} \frac{(\bar{\bar{\mu}}^{-1}_{pqs} \pm \bar{\bar{\mu}}^{-1}_{qps})(\varsigma_{ijp,mnq,s} \pm \varsigma_{mnp,ijq,s})}{1+\delta_{pq}} \det \mathsf{J}_s \ell(\boldsymbol{\xi}) d\mathcal{K}_r \tag{5.35}$$

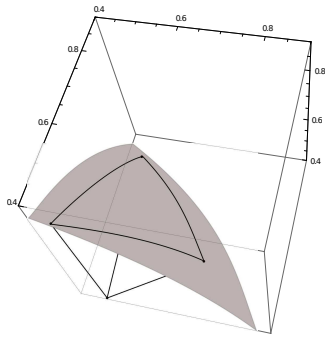
In (5.35), the bound $ij < mn$ reflects the cases where $3^1 i + 3^0 j < 3^1 m + 3^0 n$. Finally, evaluation of $\frac{1}{2}(\mathsf{S}_{uv} \pm \mathsf{S}_{vu})$ reduces to the one time evaluation of $\overset{\circ}{\mathsf{S}}{}^{\pm}_{uijvmns}$ and the element-wise evaluation of the coefficients $\mathsf{C}^{\pm}_{ijmns}$ as formulated in (5.37) followed by the multiply-add operations encoded in (5.38).

$$\overset{\circ}{\mathsf{S}}{}^{\pm}_{uij,vmn,s} \triangleq \int_{\mathcal{K}_r} \frac{\varkappa_{uij,vmn} \pm \varkappa_{vij,umn}}{2+2\delta_{ij,mn}} \ell_s(\boldsymbol{\xi}) d\mathcal{K}_r \tag{5.36}$$

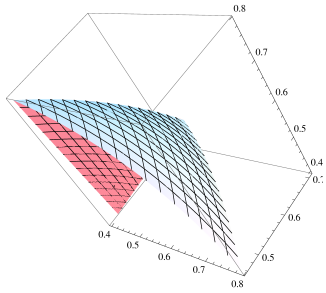
$$\mathsf{C}^{\pm}_{ij,mn,s} \triangleq \sum_{p \le q} \frac{(\bar{\bar{\mu}}^{-1}_{pq} \pm \bar{\bar{\mu}}^{-1}_{qp})(\varsigma_{ijp,mnq} \pm \varsigma_{mnp,ijq})}{1+\delta_{pq}} \det \mathsf{J} \Big|_{x=x_s} \tag{5.37}$$

$$\frac{1}{2}(\mathsf{S}_{uv} \pm \mathsf{S}_{vu}) = \sum_{\substack{i<j \\ m<n \\ ij \le mn \\ 0 \le s < S}} \mathsf{C}^{\pm}_{ij,mn,s} \overset{\circ}{\mathsf{S}}{}^{\pm}_{uij,vmn,s} \tag{5.38}$$

To evaluate the actual number of multiply-add operations required for evaluation of the stiffness matrix, one needs to observe that $\mathsf{C}^{-}_{ij,mn,s} = 0$ for cases where $ij = mn$. At the same time, for evaluation of $\mathsf{C}^{-}_{ij,mn,s}$ itself, the $(\bar{\bar{\mu}}^{-1}_{pq} - \bar{\bar{\mu}}^{-1}_{qp})$ term in (5.37) equals to zero when $p = q$. This leads to $6 \times 6 \times S$ multiply-add operations for the $+$ case and $3 \times 3 \times S$ multiply-add operations for the $-$ case. Thus, considering that these operations yield both S_{uv} and S_{vu}, the formulation requires a total of $(6 \times 6 + 3 \times 3)S/2$ multiply-add operations per matrix entry.



(a) Edge representation of the curved element.



(b) Face representation of the curved element. Only two faces are plotted here.

Figure 5.2: Construction of the F mapping from edge to face representation of a curved element.

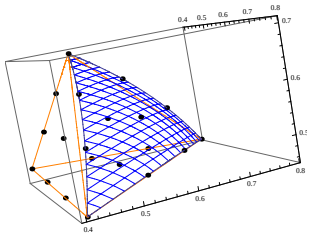


Figure 5.3: The cubic *Lagrangian* interpolation of the curved tetrahedron.

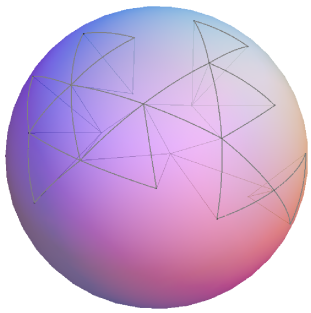


Figure 5.4: Visualization of a subset of the curved tetrahedral elements used in a mesh for the *Luneburg* lens. The surface of the lens is plotted with some transparency such that inside-sphere edges are also visible.

Notation	Definition
Γ_D	part of the boundary with essential BCs
Ω_i	i^{th} sub domain of Ω
$\bar{\bar{\mu}}_{ri}$	relative magnetic permeability tensor in Ω_i
$\mathcal{E}_i$	electrical field in sub domain i
e_i	tangent component of $\mathcal{E}_i$, i.e. $\pi_\tau(\mathcal{E}_i)$
$\hat{n}_i$	outward normal to Γ_{ij}
$\mathcal{J}^{imp}$	imposed electrical current
$\gamma_\tau(\bullet)$	trace operator $\hat{n}\wedge\bullet$
$\pi_\tau(\bullet)$	tangential trace operator $(\hat{n}\wedge\bullet)\wedge\hat{n}$
∇	del operator $\sum_i \frac{\partial}{\partial x_i}\hat{x}_i$
∇_τ	tangent component of the del operator
$\nabla\wedge\bullet$	curl operator
$\nabla\cdot\bullet$	divergence operator
$\mathbf{L}^2_0(\Omega_i)$	$\{u\in\mathbf{L}^2(\Omega_i)\vert u=0 \text{ on } \Gamma_D\}$
$H^r(\Gamma_{ij})$	*Sobolev* space of regularity r
$\mathbf{H}(\text{curl},\Omega_i)$	$\{u\in\mathbf{L}^2(\Omega_i)\vert\nabla\wedge u\in\mathbf{L}^2(\Omega_i)\}$
$\mathbf{H}_0(\text{curl},\Omega_i)$	$\{u\in\mathbf{H}(\text{curl},\Omega_i)\vert\gamma_\tau(u)=0 \text{ on } \Gamma_D\}$
$H^{-1/2}(\Gamma_{ij})$	dual space to $H^{1/2}(\Gamma_{ij})$

Table 5.1: Conformal-DDM Related Notation

5.3.5 Complexity Comparison

A brief complexity comparison between the presented approach and numerical cubature based FE matrix evaluation is presented here. Note that FE matrices are not always symmetric specially when anisotropic materials are involved. Also, it is worth mentioning that compared to the conventional UM method, the UA approach has a lower complexity mainly because of the simultaneous symmetric evaluation of uv and vu entries. However, for the sake of comparison, we assume that one uses our symmetric evaluation formulation for the numerical cubature based method too. Hence both methods at comparison will spend equal resources calculating the $\mathsf{C}^{\pm}_{\cdots}$ coefficients. Referring to (5.18), the number of multiply-add operations required for a single entry of T_{uv} equals to $\frac{6+3}{2}S(p_{geom})$ for the UA approach and $\frac{6+3}{2}\times\mathcal{G}(2n+p_{geom})\bar{\mathcal{P}}(2n+p_{geom})$ for a numerical cubature based method where $\mathcal{G}(p)$ is the number of cubature points exact for a p^{th}-order polynomial over a tetrahedral domain and where $\bar{\mathcal{P}}(p)$ denotes the average cost of evaluating an order p polynomial in $\mathbb{R}^3$ (proportional to p^3). Also, note that $S=S(p)$ is the number of points required for representation of element geometry and/or CIMPT variations, which equals

to $\frac{(p+1)(p+2)(p+3)}{6}$ for tetrahedral elements. Here we refer to Table 5.2 from [94] for typical counts of cubature points over a tetrahedral element.

p	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\mathcal{G}(p)$	1	4	8	14	14	24	36	46	61	81	109	140	171	236

Table 5.2: Number of cubature points required for accurate cubature on tetrahedra as a function of integrand polynomial order p.

As evident from the analysis above, the UA approach is clearly more efficient than the numerical cubature based method.

5.3.6 Curvature

The UA approach presented here is intrinsically capable of handling elements with curved geometry. Also, our experiments show that the method can be successfully used for evaluation of curved FE matrices such as (iso)-parametric FEs and *Non-Uniform Rational B-Spline (NURBS)* enhanced FEs [95] (see figure 5.4 and figure 5.9). The key question, however, is how to build the required bijective coordinate transformation $\digamma : \mathcal{K}_r \rightarrow \mathcal{K}_p$. Often, as it is the case with (iso)-parametric elements, a Lagrangian interpolation of an appropriate set of nodal position samples is used for $\digamma$. For the case of the spherical Luneburg lens it is easy to identify the tetrahedra that share faces and/or edges with the curved surface of the lens. For such elements, an analytical expression for the curved and rectilinear edges can be written in terms barycentric coordinates as it is shown in figure 5.2a. The edge expressions are then used to build a triangular transfinite interpolation [96] of the faces as depicted in figure 5.2b. The resulting face expressions are then used to build a tetrahedral transfinite interpolation [96] of the physical volume $\mathcal{K}_p$. Hence, the final transfinite mapping conforms exactly to the surface of the sphere. As depicted in figure 5.3, the resulting coordinate mapping is then fitted into a Lagrangian interpolation on $\mathcal{K}_r$ (3^{rd}-order in this case) and the resulting polynomial mapping is used as the actual coordinate transformation $\digamma : \mathcal{K}_r \rightarrow \mathcal{K}_p$. Nevertheless, the method discussed here can be used for arbitrary curvature as far as it is possible to have analytical expressions for the curved edges (or faces). It is worth

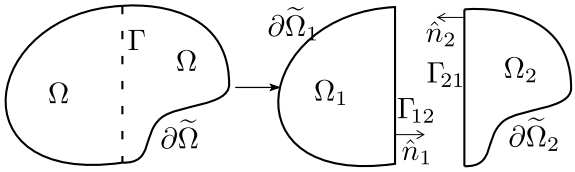


Figure 5.5: Schematic demonstration of the DDM concept

mentioning that such curvature information is usually available in terms of NURBS in the CAD model and can be made available to the FE engine at the element-matrix evaluation level.

5.4 Conformal-DDM Formulation

Non-overlapping *domain decomposition (DD)* method is an FEM-based domain decomposition that provides an efficient iterative algorithm for the FEM solution of the time-harmonic electromagnetic wave problems [89, 97–99]. The required computational resources in general are modest since only factorizations of sub-domain matrices are required. In this work, the conformal version of a non-overlapping DDM has been used in which sub-domain meshes are required to conform to each other on sub-domain interfaces. Detailed discussions on Conformal-DDM can be found in reference [89]. Herein, we only discuss the particular type/formulation of Conformal-DDM that has been used in our implementation. Note that the required notation for the following DDM related derivation is briefly listed in Table 5.1.

Referring to figure 5.5 and without loss of generality, we consider a smooth domain Ω and its decomposition into $N_p = 2$ sub-domains such that $\Omega = \Omega_1 \cup \Omega_2$. We also define two equivalent BVP one for the non-decomposed problem and another for the decomposed problem of figure 5.5. The mentioned non-decomposed and decomposed BVPs are respectively formulated in (5.39) and (5.40) through (5.43).

$$\nabla \wedge \bar{\bar{\mu}}_r^{-1} \nabla \wedge \mathcal{E} - k_0^2 \bar{\bar{\varepsilon}}_r \mathcal{E} = -\jmath k_0 \eta_0 \mathcal{J}^{\text{imp}} \qquad \text{in } \Omega \tag{5.39}$$

$$\nabla\wedge\bar{\bar{\mu}}_{r1}^{-1}\nabla\wedge\mathcal{E}_1 - k_0^2\bar{\bar{\varepsilon}}_{r1}\mathcal{E}_1 = -\jmath k_0\eta_0\mathcal{J}_1^{\text{imp}} \qquad \text{in } \Omega_1 \tag{5.40}$$

$$\nabla\wedge\bar{\bar{\mu}}_{r2}^{-1}\nabla\wedge\mathcal{E}_2 - k_0^2\bar{\bar{\varepsilon}}_{r2}\mathcal{E}_2 = -\jmath k_0\eta_0\mathcal{J}_2^{\text{imp}} \qquad \text{in } \Omega_2 \tag{5.41}$$

$$\pi_\tau(\mathcal{E}_1) = \pi_\tau(\mathcal{E}_2) \qquad \text{on } \Gamma \tag{5.42}$$

$$\gamma_\tau\left(\bar{\bar{\mu}}_{r1}^{-1}\nabla\wedge\mathcal{E}_1\right) = -\gamma_\tau\left(\bar{\bar{\mu}}_{r2}^{-1}\nabla\wedge\mathcal{E}_2\right) \qquad \text{on } \Gamma \tag{5.43}$$

In (5.40) through (5.43), $\mathcal{E}_i \in \mathbf{H}_0(\text{curl}, \Omega_i)$ represents the electric field in the i^{th} sub-domain. Note that $\mathcal{E}_1$ and $\mathcal{E}_2$ are particularly allowed to be discontinuous across Γ while equations (5.42) and (5.43) enforce the necessary tangential continuity requirements for the electric and the magnetic fields across Γ. Consequently, it can be shown that the BVP of (5.40) through (5.43) is identical to that of the original non-decomposed problem; see [89] for details.

As a matter of fact, *transmission condition (TCs)* play a critical role in the convergence of DDMs. Most recently, a few higher order TCs have been proposed that improve the numerical robustness of DDMs. Here, we used the true *second order tnransmission condition (SOTC)* . The effectiveness of SOTC for conformal DDM has been demonstrated in [89]. A general SOTC reads like (5.44) on Γ_{12} and (5.45) on Γ_{21}. Note that each 'side' of the interface is treated separately and denoted by $\Gamma_{ij}, \forall i, j \in \{1, 2\}, i \neq j$. In other words, Γ_{ij} is the interface that separates Ω_i and Ω_j and lies on $\partial\Omega_i$.

$$\begin{aligned}
&+\gamma_\tau\left(\bar{\bar{\mu}}_{r1}^{-1}\nabla\wedge\mathcal{E}_1\right)+\alpha e_1+\beta\nabla_\tau\wedge\nabla_\tau\wedge e_1\\
&+\gamma\nabla_\tau\nabla_\tau\cdot\gamma_\tau\left(\bar{\bar{\mu}}_{r1}^{-1}\nabla\wedge\mathcal{E}_1\right)\\
&=\\
&-\gamma_\tau\left(\bar{\bar{\mu}}_{r2}^{-1}\nabla\wedge\mathcal{E}_2\right)+\alpha e_2+\beta\nabla_\tau\wedge\nabla_\tau\wedge e_2\\
&-\gamma\nabla_\tau\nabla_\tau\cdot\gamma_\tau\left(\bar{\bar{\mu}}_{r2}^{-1}\nabla\wedge\mathcal{E}_2\right) \qquad \text{on } \Gamma_{12}
\end{aligned} \tag{5.44}$$

$$
\begin{aligned}
&+\gamma_\tau\left(\bar{\bar{\mu}}_{r2}^{-1}\nabla\wedge\mathcal{E}_2\right)+\alpha e_2+\beta\nabla_\tau\wedge\nabla_\tau\wedge e_2\\
&+\gamma\nabla_\tau\nabla_\tau\cdot\gamma_\tau\left(\bar{\bar{\mu}}_{r2}^{-1}\nabla\wedge\mathcal{E}_2\right)\\
&=\\
&-\gamma_\tau\left(\bar{\bar{\mu}}_{r1}^{-1}\nabla\wedge\mathcal{E}_1\right)+\alpha e_1+\beta\nabla_\tau\wedge\nabla_\tau\wedge e_1\\
&-\gamma\nabla_\tau\nabla_\tau\cdot\gamma_\tau\left(\bar{\bar{\mu}}_{r1}^{-1}\nabla\wedge\mathcal{E}_1\right)\qquad\text{on }\Gamma_{21}
\end{aligned}
\tag{5.45}
$$

In (5.44) and (5.45), the complex coefficients α, β and γ are chosen as proposed in [99] and [89]. Furthermore, on the interface Γ_{ij}, auxiliary variables $j_i \in \mathbf{L}_0^2(\Gamma_{ij})$ and ρ_i are introduced as (5.46) and (5.47).

$$
j_i=\frac{1}{k_0}\gamma_\tau\left(\bar{\bar{\mu}}_{ri}^{-1}\nabla\wedge\mathcal{E}_i\right)\qquad\text{on }\Gamma_{ij}
\tag{5.46}
$$

$$
\rho_i=\frac{1}{k_0}\nabla_\tau\cdot j_i=\frac{1}{k_0^2}\nabla_\tau\cdot\gamma_\tau\left(\bar{\bar{\mu}}_{ri}^{-1}\nabla\wedge\mathcal{E}_i\right)\qquad\text{on }\Gamma_{ij}
\tag{5.47}
$$

The introduction of ρ_i helps with convenient implementation of the $\nabla_\tau\nabla_\tau\cdot$ term in the TC. Restricted to conformal sub-domain meshes, the scalar ρ_i must be defined such that $\rho_i \in H_0^{-1/2}(\Gamma_{ij})$. With the aid of the above-mentioned auxiliary variables, TCs of (5.44) and (5.45) can be reformulated as (5.48) and (5.49) that form the actual TCs used in our implementation.

$$
\begin{aligned}
&k_0 j_1+\alpha e_1+\beta\nabla_\tau\wedge\nabla_\tau\wedge e_1+\gamma k_0^2\nabla_\tau\rho_1\\
&\quad=-k_0 j_2+\alpha e_2+\beta\nabla_\tau\wedge\nabla_\tau\wedge e_2-\gamma k_0^2\nabla_\tau\rho_2\qquad\text{on }\Gamma_{12}
\end{aligned}
\tag{5.48}
$$

$$
\begin{aligned}
&k_0 j_2+\alpha e_2+\beta\nabla_\tau\wedge\nabla_\tau\wedge e_2+\gamma k_0^2\nabla_\tau\rho_2\\
&\quad=-k_0 j_1+\alpha e_1+\beta\nabla_\tau\wedge\nabla_\tau\wedge e_1-\gamma k_0^2\nabla_\tau\rho_1\qquad\text{on }\Gamma_{21}
\end{aligned}
\tag{5.49}
$$

5.5 CIMPT Approach for the PML

Accurate realization of conformal PML layer conditions leads to artificial MPTs that are continuous functions of global coordinates and often involve curved element geometries. As a direct consequence, the CIMPT approach will be very profitable for the realization of such wave absorbing media.

Here we consider a spherical shell PML wrapped around a spherical Luneburg lens. MPTs for the conformal PML medium are obtained by means of analytical continuation of the coordinate variable that is normal to the mesh truncation surface and extends into the complex variable space. This is achieved through field transformations [90]. Here, the expressions for the PML MPTs are briefly provided. The anisotropic MPTs in a PML can be expressed as (5.50) and (5.51) where ε_0 and μ_0 respectively denote the free space permittivity and permeability constants and where $\bar{\bar{\Lambda}}$ denotes a relative permeability (permittivity) tensor.

$$\bar{\bar{\mu}} = \mu_0 \bar{\bar{\Lambda}} \tag{5.50}$$

$$\bar{\bar{\varepsilon}} = \varepsilon_0 \bar{\bar{\Lambda}} \tag{5.51}$$

In a 3D spherical coordinates system (r, θ, ϕ), $\bar{\bar{\Lambda}}$ can be written as (5.52) in which $s_r(r)$ is defined as in (5.53).

$$\bar{\bar{\Lambda}}_{r,\theta,\phi} = \hat{r}\left(\frac{\tilde{r}}{r}\right)^2 \frac{1}{s_r(r)}\hat{r} + \hat{\phi} s_r(r)\hat{\phi} + \hat{\theta} s_r(r)\hat{\theta} \tag{5.52}$$

$$s_r(r) = \alpha(r) + \jmath\frac{\sigma(r)}{\omega} \tag{5.53}$$

$$r = R_{in} + \zeta_3 \tag{5.54}$$

In (5.52), $\tilde{r}$ is a variable obtained by means of complex coordinate stretching as $\tilde{r} = R_{in} + \tilde{\zeta}_3$ where $\tilde{\zeta}_3$ is defined as in (5.55). In (5.53), ω and $\sigma(r)$ respectively denote the angular frequency and the conductivity of the PML while in (5.54) R_{in} denotes the inner radius of the spherical

PML layer and ζ_3 is the difference between r, and R_{in} signifying the normal distance form the inner surface of the PML. Finally, in (5.55) $\sigma(\kappa)$ and $\alpha(\kappa)$ are analytic functions of κ [100].

$$\tilde{\zeta_3} = \int_0^{\zeta_3} s(\kappa) d\kappa = \int_0^{\zeta_3} \left(\alpha(\kappa) + \jmath \frac{\sigma(\kappa)}{\omega} \right) d\kappa \tag{5.55}$$

In (5.52), $\bar{\bar{\Lambda}}$ is expressed in spherical coordinates and must be transformed into the Cartesian coordinate system for computer implementation. Obviously the resulting MPTs would be continuously fluctuating functions of Cartesian coordinate variables. These MPTs can be effectively handled using the CIMPT approach. Some numerical results concerning the proposed CIMPT PML are presented in section 5.6.

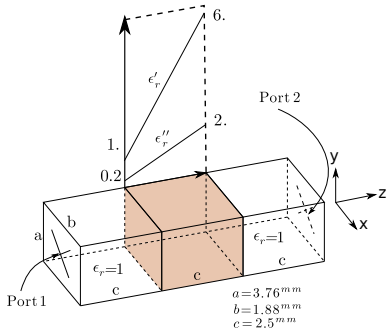
5.6 Numerical Results

This section presents the numerical results obtained using the CIMPT enabled Conformal DDM approach. The tetrahedral FE mesh used in all examples were generated using the *Cubit*[1] meshing software. The waveguide problem and all HFSS results are obtained on a laptop computer with $8\ GB$ of RAM and an Intel® Core™$i7$ CPU. Larger examples, e.g. the large Luneburg lens, are solved on a workstation computer with $48\ GB$ of RAM and two Quad-Core™ Xeon® CPUs. The Krylov subspace solver's iterations were truncated at a relative tolerance of 10^{-4}.

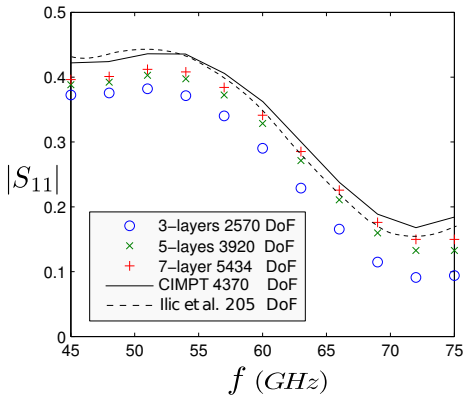
5.6.1 Waveguide

To validate the consistency of the results obtained using the UA approach, we begin with the solution of the WR-15 waveguide problem discussed in [2] which involves calculating the $|S_{11}|$ for a waveguide segment loaded by a spatially varying lossy dielectric. Both the definition and the $|S_{11}|$ solution of the waveguide problem are depicted in figure 5.6. The problem is solved using two methods: 1) a multilayer piecewise constant material model 2) a CIMPT model handled by

[1] http://cubit.sandia.gov



(a) Problem geometry.



(b) $|S_{11}|$ versus frequency.

Figure 5.6: $|S_{11}|$ versus frequency for a waveguide section with a lossy dielectric loading of $\epsilon_r := \epsilon_r' - \jmath\epsilon_r''$. The problem is identical to that of [2].

the UA method.

Method	CIMPT	3-Layer	5-layer	7-layer	Ilic et al.
DoF	4370	2570	3290	5434	205
Elements	864Tet	504Tet	648Tet	1067Tet	3Hex
CPU Time	3^{sec}	1^{sec}	2^{sec}	3^{sec}	N.A.
Field Basis	2	2	2	2	$(4,2,4)$, $(4,2,7)$
Mat. Basis	2	2	2	2	1

Table 5.3: Solution statistics for the waveguide problem. The "2" in the "Field Basis" row indicates the use of a second order curl conforming basis, i.e. the curl of the basis is complete to the 2^{nd} order. At the same time, as reported by [2] et al., (m,n,p) represents the orders of the field basis along $(\hat{x},\hat{y},\hat{z})$ directions in hexahedral elements (see figure 5.6). By material (Mat.) basis we refer to the scalar basis used for the interpolation of the MPT.

It is evident from figure 5.6 that the solution of the multilayer model approaches to that of the CIMPT model. Using the conventional method, one needs many layers to obtain results comparable to those of the CIMPT approach. On the other hand when material properties are allowed to vary, say in this case as 1^{st}-order polynomials, the actual field solution will be more complex compared to the piece-wise constant material case. Thus, field basis with higher orders of polynomial interpolation will be needed before one can fully enjoy the Degree(s) of Freedom (DoF) savings associated with the CIMPT approach. This is more evident by looking at the statistics provided in Table 5.3 where similar mesh dimensions (around $\lambda_0/4$) and same basis orders are used for all cases. Also, note that the Degree(s) of Freedom (DoF) numbers reported

(a) Electrical field intensity.

(b) Conformal-DDM's Partition of the Mesh.

Figure 5.7: Conformal-DDM simulation of the plane wave focusing effect of the *Luneburg* lens. $R = 6\lambda_0$ and $f = 10\ GHz$. The incident wave is expressed as $\mathcal{E}^i = \hat{y}e^{j\vec{k_0}\cdot\vec{r}}$ with $k_0 = |k_0|\hat{z}$. In the plot, the lens is graphically separated from its surrounding air-box to aid the viewer in seeing the surface field.

by [2] are obtained by means of an optimal choice of the basis orders along various directions. In other words, [2] exploits the fact that the actual field solution for the waveguide problem has little or no variation along the transverse directions. This, allows for a strongly anisotropic choice of the basis orders along different directions. Hence, anisotropic orders of $(4, 2, 4)$ and $(4, 2, 7)$ are used in the air and dielectric elements respectively where (m, n, p) refers to the orders of the field basis along $(\hat{x}, \hat{y}, \hat{z})$ directions.

5.6.2 Luneburg Lens

The Luneburg lens problem discussed here is a dielectric sphere with a relative permittivity tensor of $\bar{\bar{\varepsilon}} = (2 - \frac{r^2}{R^2})\bar{\bar{\mathrm{I}}}$ where r is the distance from the center of the sphere, R is the actual radius and $\bar{\bar{\mathrm{I}}}$ is the identity tensor in $\mathbb{R}^3$. Besides Luneburg's original development on the optical lens, there is plenty of work on a variety of Luneburg lens designs used in antenna/reflector devices operated in microwave and millimeter wave frequencies [101–104]. Hence, the Luneburg lens is a good example for the purpose of verifying the fidelity of the CIMPT approach.

Conceptually speaking, the Luneburg lens focuses any incoming plane wave into a point on the surface of the lens. This can be exploited to build highly directive antennas by essentially placing a receiving/transmitting element at the focus of a desired direction on the surface of the

lens. The focusing effects can be clearly observed in what will be presented in the proceeding sections. This is helpful as a means for qualitative validation of the results, e.g. figure 5.7a.

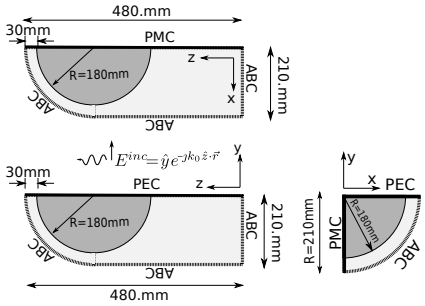


Figure 5.8: Schematic diagram showing the configuration of the quarter-*Luneburg* lens scattering problem. Curved elements were applied to the surface of the lens only.

5.6.2.1 Scattering from an $R = 6\lambda_0$ Lens

The example solved here is a $6\lambda_0$ (λ_0 is the free space wavelength) radius lens exposed to an incoming plane wave at $10\ GHz$. Figure 5.8 provides a schematic diagram of the problem geometry. In this case, the problem symmetry is exploited and only a quarter of the lens is used. The problem mesh and geometry are created using the *Cubit* mesher with a maximum discretization size of $\lambda/2.5$. Figure 5.7 shows the associated DDM partition of the FEM mesh as well as a plot of the resulting electrical field intensity. The plane-wave focusing effect can be clearly seen in figure 5.7a. The $64,592$ tetrahedron ($723,070$ Degree(s) of Freedom (DoF)) mesh was automatically partitioned into 100 sub-domains with the aid of *Metis* graph partitioning library[2]. Sub-domain matrix assemblage took 28 seconds with the memory usage peaking to $1.8\ GB$. The preconditioner was assembled in 33 seconds and demanded $3.3\ GB$ of memory (No disk caching was used) while the final solution was obtained in 88 iterations (13 seconds).

5.6.2.2 Further Comparisons for the $R = 1\lambda_0$ Lens

Also for the purpose of numerical validation, a lens with the radius of $R = 1\lambda_0$ (λ_0 is the free space wavelength) is excited by a plane wave incident from $\theta = 0$ along $\hat{z}$ and the resulting

[2] http://glaros.dtc.umn.edu/gkhome/views/metis

scattering patterns are computed using three models:

a) The continuous material index lens using the CIMPT approach. The CIMPT and curved elements are handled through the UA method.

b) Element-wise constant approximation of the Luneburg lens MPT. For each individual element, the MPTs are approximated by constant tensors which are equal to the actual MPTs at the center of the element.

c) The equivalent graded (3-layers) lens solved using the commercial HFSS solver.

Table 5.4 lists the statistics associated with this experiment. The resulting scattering patterns plotted in figure 5.9 are in reasonable agreement.

Method	CIMPT	HFSS 3-Layer	Element-wise
DoF	38323	74006	518320
Elements	5809	11417	80158
CPU Time	35^{sec}	30^{sec}	5^{min}
Field Basis	2	1	2
Mat. Basis	5	N.A. (0)	N.A. (0)
Mesh Size	$\lambda/2.5$	N.A. (h-adap.)	$\lambda/6$

Table 5.4: Solution statistics for the $R = 1\lambda_0$ *Luneburg* lens scattering problem

Solution statistics for the $R = 1\lambda_0$ *Luneburg* (see figure 5.9) lens scattering problem. The "1" and "2" in the "Field Basis" row indicate the use of first and second order curl conforming basis respectively, i.e. the curl of the first and second order basis is complete to first and second order respectively. By material (Mat.) basis we refer to the scalar basis used for the interpolation of the MPT. All data reported in this table are based on simulations performed on a laptop computer with $8\ GB$ of RAM and an Intel® Core™$i7$ CPU.

5.6.2.3 The Lens Antenna

Our next experiment involves exciting lens antennas of various radius and calculating their far field patterns. By reciprocity, it is understood that a point source on the surface of the sphere should excite a plane wave radiating from the lens. various lenses of radius $R \in \{1, 2, 3, 4, 5, 6\}\lambda_0$ are excited by a WR-90 waveguide at the frequency of $10\ GHz$ in the TE^z_{10}. As depicted in figure 5.10, the aperture of the waveguide is located 5^{mm} apart from the outer surface of the

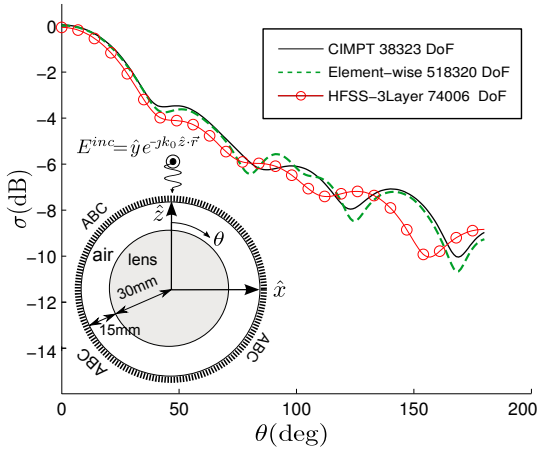


Figure 5.9: The normalized echo area σ for the $R = 1\lambda_0$ *Luneburg* lens illuminated by a plane wave plotted as a function of θ at $\phi = 0$.

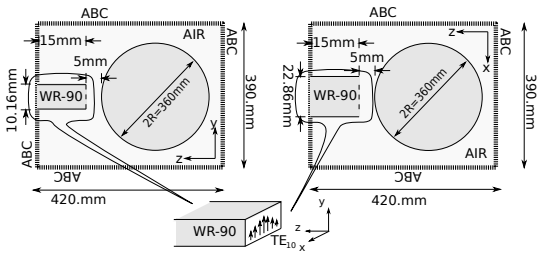


Figure 5.10: Schematic diagram showing the configuration of the $R = 6\lambda_0$ *Luneburg* lens antenna problem. Curved elements were applied to the surface of the lens only.

$R = 6\lambda_0$ lens. The resulting near field solution for the $R = 6\lambda_0$ case is plotted in figure 5.11. One can observe the flattening of the phase front as the outgoing waves propagate away from the aperture of the waveguide.

Using DDM, relatively modest computational resources are needed for computation of relatively large problems. Here, we simulated Luneburg lens antennas with lens diameters of up to $12\lambda_0$. For the $R = 6\lambda_0$ case, a mesh size of $\lambda/2.5$ results in a discretization with $423,133$ tetrahedra and $4,790,122$ Degree(s) of Freedom (DoF) (no symmetry was exploited here). Unlike the quarter lens, the *absorbing boundary condition (ABC)* surface is placed on the surface of a surrounding cube (figure 5.10). The DDM solution involves 500 automatically generated partitions . Assembling sub-domain matrices and excitation vectors took 5 minutes with a peak memory usage of $4.4\ GB$. Six minutes were spent on the construction of the preconditioner with a peak memory usage of $18\ GB$ (No disk caching was used). Note that the preconditioner's peak mem-

ory usage is directly associated with the size of the largest partition obtained by *Metis*. Hence, the number of DDM partitions was increased to keep the sub-domain dimensions more or less comparable to those of section 5.6.2.1. It took 80 iterations to finish the solution process amounting to one and a half minute of wall clock time. Figure 5.12 plots the resulting radiation patterns for the mentioned set of lens antennas with radius of $R \in \{\lambda_0, 2\lambda_0, 3\lambda_0, 4\lambda_0, 5\lambda_0, 6\lambda_0\}$.

5.6.3 Conformal PML using CIMPT

As a last example, we report the implementation of a conformal PML by means of the CIMPT approach. Figure 5.13 includes a schematic drawing of the problem geometry. Figure 5.14 shows snapshots of the field solution of a $1\lambda_0$ radius Luneburg lens antenna wrapped in a continuous PML. The $0.25\lambda_0$ thick PML is placed $0.25\lambda_0$ away from the surface of the lens. The attenuation function of the PML is set to $s_r = 1 + \jmath\sigma_{max}\zeta_3/2\pi\varepsilon_0$ in which ζ_3 is the distance from any point within the PML region to the inner surface of the PML layer. Here we use $\sigma_{max}=1$. Similar to the $R=1\lambda_0$ lens discussed in section 5.6.2.2, the geometry is meshed by *Cubit* using a maximum discretization size of $\lambda/2.5$. The PML exhibits strong attenuation of the radiated field which validates its proper implementation. Figure 5.15 plots the resulting radiation pattern in the $\phi=0$ plane. The observed CPU time and memory usage are pretty much the same as those reported for the $R=1\lambda_0$ lens discussed in section 5.6.2.2.

5.7 Conclusion

An efficient UA approach capable of handling curved geometries and continuously varying MPTs was introduced. Compared to the conventional approach for evaluation of FE matrices, the presented UA approach assumes that material properties and element Jacobian related terms can vary as continuous functions across the element. With the aid of appropriate interpolation schemes the required cross-element continuity or discontinuity of the desired terms can be guaranteed. This leads to more flexible and accurate modeling of the underlying physics.

Merged with a conformal DDM-FEM solver, the UA approach was validated on a few problems including Luneburg lenses and a spherical perfectly matched layer for which the MPTs were again modeled as continuous functions of spatial coordinates. The non-overlapping conformal DDM used in this work is based on SOTCs. The true SOTC which enforces the tangential continuity of the fields provides an efficient iterative algorithm for the FE solution of time-harmonic Maxwell's equations. Finally, relatively large electrical problems with continuously changing material properties were solved and presented for the purpose of numerical validation.

Despite the current verification on the Luneburg lens problem or the conformal PML, the CIMPT technique is believed to be useful in other areas such as: a) multi-physical problems where dielectric properties change due to other physical phenomena while re-meshing of the problem domain needs to be avoided, e.g. dielectric permittivity may change due to temperature/pressure fluctuations; b) in doped semiconductors and plasma antennas/reflectors, dielectric properties may follow a continuous profile following various physical phenomena. Hence, the introduction of the UA method will play an influential role in future developments associated to the CIMPT approach. Finally, it should be stated here that the current implementation of our code uses second order curl conforming basis functions. Last but not least, it worth mentioning that complete utilization of the benefits of the proposed CIMPT and UA methods can be achieved by means of implementations with higher orders of field basis functions.

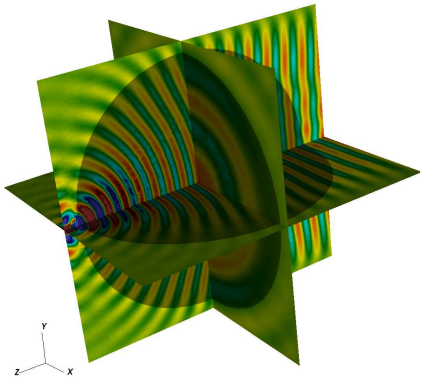


Figure 5.11: A three-slice field plot of the $\mathcal{E} \cdot \hat{y}$ field component in the continuous index *Luneburg* lens antenna with $R = 6\lambda_0$ and $f = 10\,GHz$. The shaded part of the plot belongs to the dielectric lens.

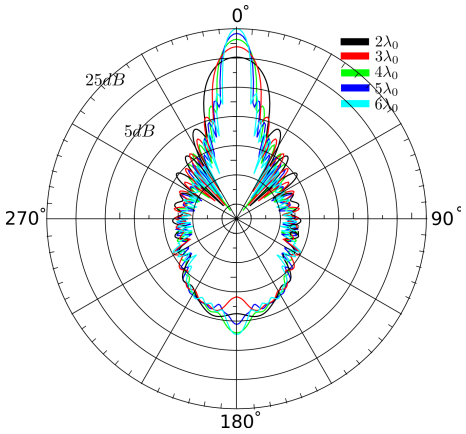


Figure 5.12: The radiation pattern (dB) of *Luneburg* lens antennas of various radius R as function of θ at $\phi = 0$.

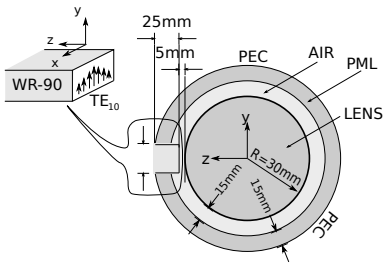
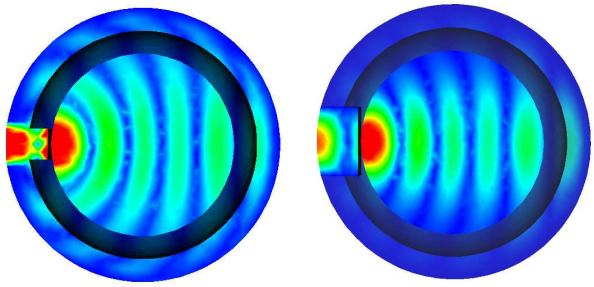


Figure 5.13: Schematic diagram showing the configuration of the PML-lens antenna problem.



(a) yz-plane. (b) zx-plane.

Figure 5.14: Snapshots of the magnitude of the electric field distribution in the antenna and the PML region.

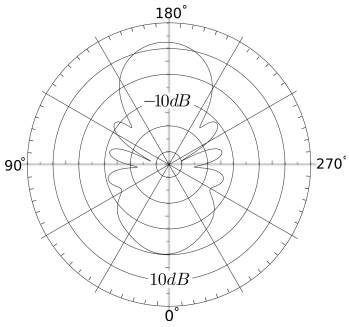


Figure 5.15: The radiation pattern (dB) of the *Luneburg* lens antenna with $R = \lambda_0$ at $\phi = 0$

Bibliography

[1] Davood Ansari Oghol Beig and Mook Seng Leong. Dual-grid based tree/cotree decomposition of higher-order interpolatory $H(\nabla\wedge,\Omega)$ basis. *Int. J. Numer. Meth. Engng.*, 2010.

[2] Milan M. Ilic, Andjelija Z. Ilic, and Branislav M. Notaros. Continuously inhomogeneous higher order finite elements for 3-D electromagnetic analysis. *IEEE Trans. Ant. Prop.*, 57 (9), September 2009.

[3] J.T. Oden. Finite element methods part 1. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis Vol. 2*. Elsevier Science, North-Holland, 2003.

[4] O.J. Antunes, J.P.A. Bastos, and N. Sadowski. Using high-order finite elements in problems with movement. *IEEE Trans. Magn.*, 40(2), March 2004.

[5] Roberto D. Graglia and Guido Lombardi. Singular higher order complete vector bases for finite methods. *IEEE Trans. Ant. Prop.*, 40(2), March 2004.

[6] J.S. Hesthaven and T. Warburton. Nodal high-order methods on unstructured grids, i. time-domain solution of maxwells equations. *J. Comput. Phys.*, 181:186–221, 2002.

[7] L.E. Garcia-Castillo, A.J. Ruiz-Genoves, I. Gomez-Revuelto, M. Salazar-Palma, and T.K. Sarkar. Third-order nedelec curl-conforming finite element. *IEEE Trans. Magn.*, 38(5), September 2002.

[8] Wei Cai, Tiejun Yu, Han Wang, Yijun Yu, et al. High-order mixed rwg basis functions for electromagnetic applications. *IEEE Trans.Microw. Theory and Tech.*, 49(7), 2001.

[9] J.P. Webb. Hierarchal vector basis functions of arbitrary order for triangular and tetrahedral finite elements. *IEEE Trans. Ant. Prop.*, 47(8), August 1999.

[10] Milan M. Ilic, Andjelija Z. Ilic, and Branislav M. Notaros. Higher order large-domain FEM modeling of 3-D multiport waveguide structures with arbitrary discontinuities. *IEEE Trans. Microw. Theory and Tech.*, 52(6), June 2004.

[11] Roberto D. Graglia and Ioan-L. Gheorma. Higher order interpolatory vector bases on pyramidal elements. *IEEE Trans. Ant. Prop.*, 47(5), October 1999.

[12] Roberto D. Graglia, Donald R. Wilton, and Andrew F. Peterson. Higher order interpolatory vector bases for computational electromagnetics. *IEEE Trans. Ant. Prop.*, 45(3):329–342, March 1997.

[13] Milan M. Ilic and Branislav M. Notaros. Higher order hierarchical curved hexahedral vector finite elements for electromagnetic modeling. *IEEE Trans. Microw. Theory and Tech.*, 51(3), March 2003.

[14] Pavel Solin, Karel Segeth, and Ivo Delezel. *Higher Order Finite Elements*. Chapman & Hall/CRC, 2004.

[15] Pavel Solin and Karel Segeth. Hierarchic higher-order hermite elements on hybrid triangular/quadrilateral meshes. *Math. Comput. Simulat.*, 76(1-3):198–204, October 2007. ISSN 0378-4754.

[16] Sebastien Jund and Stephanie Salmon. Arbitrary high-order finite element scemes and high-order mass lumping. *Int. J. Appl. Math. Comput. Sci.*, 17(3):375–393, 2007.

[17] M.A. Taylor and B.A. Wingate. A generalized diagonal mass matrix spectral element method for non-quadrilateral lements. *Appl. Numer. Math.*, 33(1):259–265, May 2000.

[18] Richard Pasquetti and Francesca Rapetti. Spectral element methods on triangles and quadrilaterals: comparisons and applications. *J. Comput. Phys.*, 198:349–363, 2004.

[19] Qi Chen and Ivo BabuSkab. Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle. *Comput. Meth. Appl. Mech. Eng.*, 128:405–417, 1995.

[20] Qi Chen and Ivo BabuSkab. The optimal symmetrical points for polynomial interpolation of real functions in the tetrahedron. *Comput. Meth. Appl. Mech. Eng.*, 137:89–94, 1996.

[21] M.A. Taylor and B.A. Wingate. Fekete collocation points for triangular spectral elements. *SIAM J. Numer. Anal.*, 1998.

[22] J.S. Hesthaven. From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex. *SIAM J. Numer. Anal.*, 35(2):655–676, 1998.

[23] L. Bos, M.A. Taylor, and B.A. Wingate. Tensor product gauss-lobatto points are fekete points for the cube. *SIAM J. Numer. Anal.*, 70(236):1543–1547, 2001.

[24] M.A. Taylor, B.A. Wingate, and R.E. Vincent. An algorithm for computing fekete points in the triangle. *SIAM J. Numer. Anal.*, 38:1707–1720, 2000.

[25] M.G. Blyth, H. Luo, and C. Pozrikidis. A comparison of interpolation grids over the triangle or the tetrahedron. *J. Eng. Math.*, 56:263–272, 2006.

[26] H. Luo and C. Pozrikidis. A lobatto interpolation grid in the tetrahedron. *IMA J. Appl. Math.*, 71:298–313, 2006.

[27] M.G. Blyth and C. Pozrikidis. A lobatto interpolation grid over the triangle. *IMA J. Appl. Math.*, 71:153–169, 2006.

[28] T. Warburton. An explicit construction of interpolation nodes on the simplex. *J. Eng. Math.*, 56:247–262, 2006.

[29] Joon-Ho Lee, Tian Xiao, and Qing H. Liu. A 3-D spectral-element method using mixed-order curl conforming vector basis functions for electromagnetic fields. *IEEE Trans. Microw. Theory and Tech.*, 54(1), January 2006.

[30] Robert N. Rieben, Daniel A. White, and Garry H. Rodrigue. Improved conditioning of finite element matrices using new high-order interpolatory bases. *IEEE Trans. Ant. Prop.*, 52(10), October 2004.

[31] P. Bettes and O. C. Zienkiewics. Diffraction and reflection of surface waves using finite and infinite elements. *Int. J. Numer. Meth. Engng.*, 11:1271–1290, 1977.

[32] P. Bettes. Infinite elements. *Int. J. Numer. Meth. Engng.*, 11:53–64, 1977.

[33] G. Beer and J. L. Meek. Infinite domian elements. *Int. J. Numer. Meth. Engng.*, 17:43–52, 1981.

[34] Paul P. Lyn. Infinite elements with $\mathbf{1}/\boldsymbol{r^n}$ type decay. *Int. J. Numer. Meth. Engng.*, 17: 347–355, 1981.

[35] David S. Burnett. A three dimensional acoustic infinite element based on a prolate spehroidal multiple expansion. *J. Acoust. Soc. Am.*, 96(5):2798–2815, 1994.

[36] Exterior problems of wave propagation. *Comput. Meth. Appl. Mech. Eng.*, 164(1-2), 1988.

[37] L. Demkowicz and M. Pal. An infinite element for maxwell equations. *Comput. Meth. Appl. Mech. Eng.*, 164(1-2), 1998.

[38] Dan Givoli and Igor Patlashenko. An optimal high order non-reflecting finite element scheme for wave scattering problems. *Int. J. Numer. Meth. Engng.*, 53:2389–2411, 2002.

[39] Yang Ruiliang and Wang Hongzhen. A novel prolate spheroid infinite element. In *Proceedings of* $\mathbf{8^{th}}$ *International Seminar/Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory*, pages 190–193, 2003.

[40] W. Cecot, L. Demkowicz, and W. Rachowicz. A two dimensional infinite element for maxwell's equations. *Comput. Meth. Appl. Mech. Eng.*, 188:625–643, 2000.

[41] W. Cecot, L. Demkowicz, and W. Rachowicz. An hp-adaptive finite element method for electromagnetics. part 3: A three-dimensional infinite element for maxwell's equations. *Int. J. Numer. Meth. Engng.*, 57:899–921, 2003.

[42] R.J. Astley and J.P. Coyette. Conditioning of infinite element schemes for wave problems. *Commun. Numer. Meth. Engng.*, 17:31–41, 2001.

[43] R.J. Astley and J.P. Coyette. The performance of spheroidal infinite elements. *Int. J. Numer. Meth. Engng.*, 52:1379–1396, 2001.

[44] D. Dreyer and O. von Estor. Improved conditioning of innite elements for exterior acoustics. *Int. J. Numer. Meth. Engng.*, 58:933–953, 2003.

[45] L. Demkowicz. *Computing with hp-Adaptive Finite Elements*, volume 1. Chapman & Hall/CRC, 2007.

[46] Edward L. Wilson. The static condensation algorithm. *Int. J. Numer. Meth. Engng.*, 8(1): 198–203, 1973.

[47] Tomas Vejchodsky and Pavel Solin. Static condensation, partial orthogonalization of basis functions, and ilu preconditioning in the hp-FEM. *J. Comput. Appl. Math.*, 218(1):192–200, 2008.

[48] M Dufresne and P P Silvester. Universal matrices for the n-dimensional finite element. *Computation in Electromagnetics, Third International Conference on (Conf. Publ. No. 420)*, pages 223–228, 1996.

[49] D. Villeneuve and J.P.Webb. Universal matrices for high order finite elements in nonlinear magnetic field problems. *IEEE Trans. Magn.*, 33(5), September 1997.

[50] Michael Heroux, Roscoe Bartlett, Vicki Howle, Robert Hoekstra, Jonathan Hu, Tamara Kolda, Richard Lehoucq, Kevin Long, Roger Pawlowski, Eric Phipps, Andrew Salinger, Heidi Thornquist, Ray Tuminaro, James Willenbring, and Alan Williams. An overview of trilinos. Technical Report SAND2003-2927, Sandia National Laboratories, 2003.

[51] J.P. Webb. Edge elements and what they can do for you. *IEEE Trans. Magn.*, 29(2), March 1993.

[52] J.B. Manges and Z.J. Cendes. Tree-cotreee decompositions for first-order complete tangential vector finite elements. *Int. J. Numer. Meth. Engng.*, 40:1667–1685, 1997.

[53] P. Alotto and I. Perugia. Mixed finite element methods and tree-cotree implicit condensation. *IEEE Trans. Magn.*, 34(4):1523–1526, July 2000.

[54] Piergiorgio Alotto and Ilaria Perugia. Tree-cotree implicit condensation in magnetostatics. *IEEE Trans. Magn.*, 36(4), July 2000.

[55] H. Igarashi. On the property of the curl–curl matrix in finite element analysis with edge elements. *IEEE Trans. Magn.*, 37(5), September 2001.

[56] Seung-Cheol Lee, J.-F. Lee, and Robert Lee. Hierarchical vector finite elements for analyzing waveguiding structures. *IEEE Trans. Microw. Theory and Tech.*, 51(8):1897–1905, August 2003.

[57] M. Repetto and F. Trevisan. Global formulation of 3D magnetostatics using flux and gauged potentials. *Int. J. Numer. Meth. Engng.*, 60:755–772, 2004.

[58] Neelakantam V. Venkatarayalu and J.-F. Lee. Removal of spurious dc modes in edge element solutions for modeling three-dimensional resonators. *IEEE Trans. Microw. Theory and Tech.*, 54(7):3019–3025.

[59] B. Doliwa, H. de Gersem, T. Weiland, and T. Boonen. Optimised electromagnetic 3D field solver for frequencies below the first resonance. *IET Sci. Meas. Technol.*, 1(1):53–56, 2007.

[60] Shih-Hao Lee and Jian-Ming Jin. Applicaiton of the tree-cotree splitting for improving matrix conditioning in the full-wave finite-element analysis of high-speed circuits. *Microw. Opt. Tech. Lett.*, 50(6), June 2008.

[61] Davood Ansari Oghol Beig. Semi-optimal interpolation nodes and improved FE matrices. *Submitted To J. Sci. Comput.*, February 2010.

[62] Davood Ansari Oghol Beig, Z. Peng, J.-F. Lee, and Jue Wang. A universal matrix approach for continous material properties in DDM. *IEEE Trans. Ant. Prop.*, May 2010.

[63] Davood Ansari Oghol Beig, Ya Qiong Zhang, Ban Leong Ooi, and Mook Seng Leong. Investigation of application of higher-order elements in FEM. In *Interna- tional Symposium on Antennas and Propagation*, 2006.

[64] Tony Low and Davood Ansari Oghol Beig. Signatures of quantum transport through two-dimensional structures with correlated and anti-correlated interfaces. *Phys. Rev. B.*, 2008.

[65] J. C. Nedelec. Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 35:315–341, 1980.

[66] J. C. Nedelec. A new family of mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 50:57–81, 1986.

[67] D. A. White and J.M. Koning. Computing solenoidal eignemodes of the vector helmholtz equation: A novel approach. *IEEE Int. AP-S Symp. Dig., Washingotn DC*, pages 193–196, July 2005.

[68] S. Perepelitsa, R. Dyczij-Edlinger, and J.-F. Lee. Finite-element analysis of arbitrarily shaped cavity resonator using $H^1(curl)$ elements. *IEEE Trans. Magn.*, 33(3):1776–1779, March 2004.

[69] Johanthan S. Golan. *Foundations of Linear Algebra*. Kluwer Academic Publishers, 1995.

[70] J. Tinsley Oden and L. Demkowicz. *Applied Functional Analysis*. CRC Press, 1996.

[71] Peter Monk. A simple proof for an edge element discretization of maxwell's equations. *Comput. Electromagn.*, 28:127–142, 2003.

[72] V. Girault and P.A. Raviart. *Lecture Notes in Mathematics Finite Element Approximation of the Navier-Stokes Equations*, volume 749. Springer-Verlag, 1979.

[73] Matthew N.O. Sadiku. *Numerical Techniques in Electromagnetics*. CRC Press, second edition edition, 2000.

[74] C. W. Crowley, P. P. Silvester, and H. Hurwitz. Covariant projection elements for 3-D vector field problems. *IEEE Trans. Magn.*, 24:397–400, January 1988.

[75] Constantine A. Balanis. *Advanced Engineering Electromagnetics*. John Wiley & Suns , Inc., 1989.

[76] A Taflove. *The Finite-Difference Time Domain Method in Computational Electrodynamics*. Artech House, Boston, 1999.

[77] A. Sommerfeld. *Partial Differential Equations in Physics*. Academic Press, New York, 1949.

[78] Atef Elsherbeni, Darko Kajfez, and Sheng Zheng. Circiular sector waveguides. *IEEE Trans. Ant. Prop.*, 33(6), 1991.

[79] Gilbert Strang and George J. Fix. *An Analysis of the Finite Element Method*. Prentice Hall, NJ, 1989.

[80] A. Safjan and M. Newmen. The ill-conditioning of infinite element stiffness matrices. *Comput. Math Appl.*, 41:1263–1291, June 2001.

[81] L. Bos. Bounding the lebesgue function for lagrange interpolation in the simplex. *J Approx. Theory*, 38:43–59, 1983.

[82] Bloom T., Bos L., C. Christensen, and N. Levenberg. Polynomial interpolation of holomorphic function in c and c^n. *Rocky Mt. J. Math.*, 22(2):441–469, 1992.

[83] Roth M.J. *Nodal Configurations and Voronoi Tessellations for Triangular Spectral Elements*. PhD thesis, University of Victoria, 2005.

[84] M.A. Taylor, B.A. Wingate, and L. Bos. Tensor product gauss-lobatto points are fekete point for the cube. *Math. Comput.*, 70:1543–1547, 2001.

[85] Dubiner M. Spectral methods on triangles and other domains. *J. Sci. Comput.*, 6(4):345–390, 1991.

[86] Jian-Ming Jin. *The Finite Element Method in Electromagnetics*. John Wiley & Sons Inc., 2002.

[87] Enrica Martini, Giuseppe Pelosi, and Stefano Selleri. With a new type of curvilinear mapping for the analysis of microwave passive devices. *IEEE Trans. Ant. Prop.*, 51(6), June 2003.

[88] J S Wang and N Ida. Curvilinear and higher order edge finite elements in electromagnetic field computation. *IEEE Trans. Magn.*, 29(2), March 1993.

[89] V. Rawat. *Finite Element Domain Decomposition with Second Order Transmission Conditions for Time-Harmonic Electromagnetic Problems*. PhD thesis, Ohio State University, 2009.

[90] F.L. Teixeira and W.C. Chew. Analytical derivation of a conformal perfectly matched absorber for electromagnetic waves. *Microw. Opt. Tech. Lett.*, 17:231–236, 1997.

[91] Elsevier Prev, D. Schurig, and D. R. Smith. Controlling electromagnetic fields. *Science*, 312(5781):1780–1782, June 2006.

[92] J. Berenger. Three-dimensional perfectly matched layer for the absorbtion of electromagnetic waves. *J. Comput. Phys.*, 127(2):363–379, 1996.

[93] Pavel Solin and Karel Segeth. *Higher-Order Finite Element Methods*. Chapman & Hall. ISBN 158488438X.

[94] Linbo Zhang, Tao Cui, and Hui Liu. A set of symmetric quadrature rules on triangles and tetrahedra. *J Comput Math*.

[95] Ruben Sevilla, Sonia Fernández-Méndez, and Antonio Huerta. Nurbs-enhanced finite element method (NEFEM). *International Journal for Numerical Methods in Engineering*, 76 (1):56–83, February 2008.

[96] Alain Perronnet. Interpolation transfinie sur le triangle, le tétraèdre et le pentaèdre. application à la création de maillages et à la condition de dirichlet. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 326(1):117–122, 1998.

[97] K. Zhao, V. Rawat, and J.-F. Lee. A domain decomposition method with non-conformal meshes for finite periodic and semi-periodic structures. *IEEE Trans. Ant. Prop.*, 55(9): 2559–2570, September 2007.

[98] Y.-J. Li and J.-M. Jin. A new dual-primal domain decomposition approach for finite element simulation of 3-D large-scale electromagnetic problems. *IEEE Trans. Ant. Prop.*, 55 (10):2803–2810, October 2007.

[99] Z. Peng, V. Rawat, and J.-F. Lee. One way domain decomposition method with second order transmission conditions for solving electromagnetic wave problems. *J. Comput. Phys.*, 229:1181–1197, 2010.

[100] S. Dosopoulos and J.-F. Lee. Interior penalty discontinuous galerkin finite element method for the time-dependent first order maxwell's equations. *IEEE Trans. Magn.*, 46(8):3512–3515, August 2010.

[101] Benjamin Fuchs, Olivier Lafond, Sébastien Palud, Laurent Le Coq, Mohamed Himdi, Michael C. Buck, and Sébastien Rondineau. Comparative design and analysis of luneburg and half maxwell fish-eye lens antennas. *IEEE Trans. Ant. Prop.*, 56(9):3058–3062, September 2008.

[102] Hossein Mosallaei and Yahya Rahmat-Samii. nonuniform luneburg and two-shell lens antennas: radiation characteristics and design optimization. *IEEE Trans. Ant. Prop.*, 49(1): 60–69, January 2001.

[103] Benjamin Fuchs, Laurent Le Coq, Olivier Lafond, Sébastien Rondineau, and Mohamed Himdi. Design optimization of multishell luneburg lenses. *IEEE Trans. Ant. Prop.*, 55(2): 283–289, January 2007.

[104] Slavi R. Baev, Boyan N. Hadjistamov, and Plamen I. Dankov. Modeling and simulations of lüneburg lens antennas for communication purposes. In *The* 16^{th} *Telecommunications Forum*, TELEFOR, pages 488–491, 2008.

Appendix A

The ‘Burn and Proceed’ Algorithm

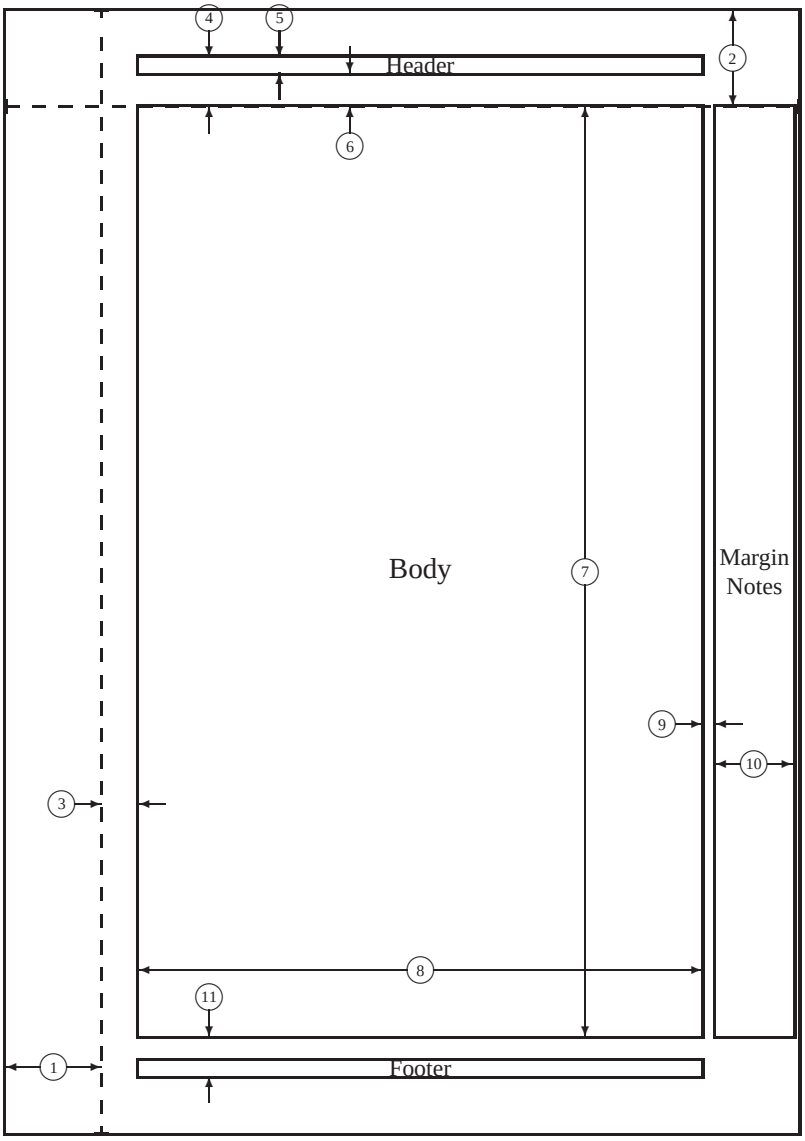
Implementation of HO spectral FE requires the realization of a higher order connectivity generator. In simple words, local and global Degree(s) of Freedom (DoF), node and element numbering must be available to the FE engine before FE global matrices can be assembled. It is out of the question that there are many ways for this to be realized. In this work, however, we start with low order mesh connectivity information and build the required higher order connectivity data as a mesh post processing step. Intuitively, the algorithm that is presented here will be called “the burn and proceed” algorithm because of the way it covers the lower order mesh which is very similar to how fire spots propagate on a ground with sparse distribution of brushwood. Note that this is an order N complexity algorithm with respect to the number of Degree(s) of Freedom (DoF). In the following, by “element list” we refer to the list that returns the global node numbers associated to each element. Furthermore, the inverse element list is a list that returns the list of elements associated to each node and the node list is the list of valid node numbers. A short pseudo-code description of the algorithm is reflected in figure A.1.

1. Populate the inverse element list.
2. Set ⓘ = 0 ;
   a. Assuming that the current node is node ⓘ, identify if ⓘ is already burned.
   b. If burned, go to 2d, otherwise, continue with the following:
      i. Through the inverse element list, find the elements associated to ⓘ and generate internal nodes for each of them.
      ii. Through the inverse element list and the element list find the nodes that from lines connecting to ⓘ.
      iii. On each line (starting from ⓘ) ending to an unburned node generate the required element boundary nodes and associate them the appropriate element(s).
      iv. Mark ⓘ as burned.
   c. If the end of node list has been reached go to 3.
   d. Increase ⓘ to the next value and go to 2.
3. End.

Figure A.1: The "burn and proceed" algorithm.

Appendix B

Document Layout



```
 1  one inch + \hoffset       2  one inch + \voffset
 3  \oddsidemargin = 28pt     4  \topmargin = -37pt
 5  \headheight = 12pt        6  \headsep = 25pt
 7  \textheight = 700pt       8  \textwidth = 424pt
 9  \marginparsep = 10pt     10  \marginparwidth = 59pt
11  \footskip = 30pt             \marginparpush = 5pt (not shown)
    \hoffset = 0pt               \voffset = 0pt
    \paperwidth = 597pt          \paperheight = 845pt
```

Glossary

$1D$ one dimensional/dimensions. 83
$2D$ two dimensional/dimensions. 51, 57, 75
$3D$ three dimensional/dimensions. 20, 50–52, 60, 75
$H(\nabla\wedge)$ curl conforming functional space. vii, 3, 4, 11, 12, 18, 51, 53, 90, 93, 101–104, 107, 108, 110, 112
H^1 H^1 functional space. vii, 3, 4, 6, 10, 11, 18, 30, 53, 75, 76, 83, 90, 92, 100, 104, 106, 111, 112
L^2 L^2 functional space. 70

affine . 23
AHO arbitrary high order. 3–5, 7, 10
Ampere . 58, 111
Anasazi The Anasazi subpackage of the Trilinos project, see http://trilinos.sandia.gov/. 53
AztecOO The AztecOO package of the Trilinos project, see http://trilinos.sandia.gov/. 53

Babuska . 91
barycentric . 17, 23
BEM boundary element method. 5
Bessel . 75, 81
bijective . 23
BKS Block Krylov Schur. 53, 54, 56
burn and proceed . 9
BVP boundary value problem. 3, 10, 125, 126

CAD computer aided design. 8
CAM computer aided manufacturing. 8
Cartesian . 51
Cauchy . 70
CGS conjugated gradient squared. 58
chain . 29
CIMPT continuously inhomogeneous material property tensor. 2–4, 9, 12, 128, 130
complement . 39, 108
condition . vi, vii, 4, 89–93, 95, 97–100, 106, 107, 112
cotree . 17–19, 32, 33, 38–41, 49, 53
CPU Central Processing Unit. 46, 60
curvilinear . 22, 23, 27, 76

DC direct current. 42
DDM domain decomposition method. vi, 1–3, 115, 116, 126
determinant . 91, 100
Dirichlet . 9, 125
DoF Degree(s) of Freedom. 7, 9, 10, 14, 18, 19, 26, 30–35, 37–43, 46, 49, 51–53, 58, 104, 112, 140

dual . vi, vii, 2, 10, 11, 14, 19, 20, 40, 41, 43–47, 49–51, 59, 60, 104

eigenproblem . 107
eigenvalue . 55
eigenvalue . 55
eigenvalue . 107, 108
electrostatic . 91
electrostatics . 98
EM electromagnetics. vi, 3, 5, 6, 9, 17, 59, 90, 102, 108
Epetra a sublibrary of the Trilions package. 9
Epetra The Epetra package of the Trilinos project, see http://trilinos.sandia.gov/. 47
Euler . 38
EVP eigenvalue problem. 3, 10
exteriority . 2, 4, 9, 12

F/IE finite/infinite element. vii, 12, 68, 69, 86, 87
FE finite element. vi, vii, 1–7, 9–14, 17, 18, 20, 23, 44, 58, 59, 66, 68–70, 76, 82, 83, 85–87, 89, 90, 92, 93, 98, 99, 102, 109, 110, 112, 130, 140
Fekete . 4, 90–92, 97–101, 106, 107
FEM finite element method. vii, 1–11, 13, 17, 18, 21, 23, 42–44, 46, 47, 54, 55, 57–59, 65, 89, 91, 92, 95, 104, 107, 109, 114
Frobenius . 95–97, 106

Galerkin . 54
GiNaC GiNaC library. 9
GLL Gauss-Lobatto-Legendre (points). 91
GLT Gauss-Lobatto-Tchebychev (points). 92
GMRES generalized minimal residual. 99
Green . 13, 69
GVD Generalized Vandermonde Determinant. 91
GVM Generalized Vandermonde Matrix. 91, 94, 97, 100, 101, 104, 106, 112

Hankel . 75
HDF5 hierarchical data format 5. 9
Helmholtz . 66, 67, 69, 81, 90
hermitian . 10
hierarchical . vi, 3, 6, 7, 10, 14, 18, 19, 59
Hilbert . 16, 21
HO higher order. vi, vii, 1–3, 5, 7, 9–14, 18–20, 22, 28, 59, 98, 108, 140
HOC higher order connectivity. 9
hypercube a geomtrical object denoted by $\square^d$. 4, 101
hypervolume . 100

IE infinite element. vi, vii, 2, 4, 5, 9, 12, 65, 66, 68, 69, 71, 76, 83–87
IEM infinite element method. 4, 5, 9
isoparametric . 82

Jacobian . 12, 16, 28, 70, 77, 115–117

Koornwinder . 94, 103
KPI Khayyam-Pascal isometric. 93, 98, 104, 106–108, 111
Kronecker . 16
Krylov . 19, 44, 53, 57, 111

Robin . 9

Salome The Salome package, see http://www.salome-platform.org/home. 8, 9
Schrodinger . 13
Schwarz . 70
separation . 108
shape . 93, 94, 103, 105
Silvester . 93
simplex a geomtrical object denoted by $\triangle^d$. 4, 50–53, 60, 89, 92
Sobolev . 18, 125
Sommerfeld . 67, 68, 70
SOTC second order transmission condition. 126
spectral . vi, vii, 2, 3, 5–7, 9–12, 14, 19, 59, 89, 108, 140
spurious . vii, 17, 18, 20, 55, 57, 108
static condenstaion . 7
steady state . 6
stiffness . 85, 93, 97, 102, 107
submultiplicative . 96, 97
subtree . 49, 50
symmetric . 10, 96, 97

TC transmission condition. 127
TC tree/cotree. vi, vii, 2, 11, 12, 14, 17, 18, 20, 26, 36, 50, 55
Tchebychev . 92
tree . 11, 17–20, 26, 32, 33, 38–41, 43–51, 53, 60
Trilinos The Trilinos package, see http://trilinos.sandia.gov/. i, 9, 10, 53

under determined . 9, 106
universal . 9, 23

Vandermonde . 12, 92
VTK The visualization toolkit, see http://www.vtk.org/. 10

warp & blend . 90–92, 107, 108, 111, 112

Summary

Deployment of HO basis functions for FE analysis of EM problems is a tempting task. For a p-order complete basis, the optimal h^{p+1} rate of convergence should translates to tremendous amounts of saving in DoF and computer resources. There are many reasons, however, why this has not been (and probably will not be) achieved in whole. a) the presence of small features in the geometry of almost all practical problems inhibits the use of larger HO elements b) the limited capability of mesh generators in producing HO curved meshes c) ill conditioning of HO FE matrices. There is an ongoing endeavor to eliminate or elevate these limitations. In the context of hierarchical FEs, this has lead to the development of hp-adaptive methods while in the context of spectral FEs the problem can also be seen from a slightly different angle. The emergence of DDM has raised the possibility of applying HO FEs onto sub-domains where at least some the abovementioned limitations can be partially elevated. In this regard, this research will be focused on the following main objectives:

1. Improving the condition numbers of HO FE and IE matrices.
2. Developing schemes for efficient evaluation of element matrices with complex geometries and/or material properties.
3. Also, in order to avoid the difficulties associated with spurious modes, FE simulation of EM problems requires proper treatment of the null-space of the curl operator. Hence, we shall introduce a new dual-grid based the T/C decomposition method for higher order spectral elements.

Item 1 is the focus of chapter 4 and part of chapter 3 where the properties of interpolation nodal sets are exploited for construction of improved FE and IE basis functions. The results are promising and indicate that condition number improvements could be as high 10^2 or 10^3 depending on the order of the polynomial basis. These condition number improvements are studied for both $H(\nabla\wedge)$ and H^1 type problems.

FE modeling of wave and scattering phenomena is made possible by means of special mesh truncation techniques. Here, the IE method was chosen because of its similarity with FEs that allowed the extension of the nodal set-based matrix conditioning technique into IEs. Again, it was shown that significant condition number improvements can be achieved while it is observed that the resulting F/IE matrix condition numbers can still grow undesirably as problem dimensions and basis orders continue to grow. This, however, is believed to be elevated by deployment of IE dedicated preconditioners and might turn into a good topic for further research.

The issue of spurious modes in FE solution of Maxwell equation must be addressed by means of an appropriate null-space treatment. This is systematically done through T/C decomposition. In chapter 2 a new dual-grid based T/C technique for HO spectral elements is introduced. The method has certain advantages over its predecessors. Most of all, it involves no global path integration operators which could be costly for HO elements. Through the process of this develop-

ment, it is also shown that the resulting constraint matrices are parley determined by the topology of the FE mesh.

One of the advantages of FEM over *method of moments (MoM)* is that the matrices can be analytically and exactly evaluated. Moreover, material properties are traditionally (and counter-intuitively) treated as element-wise constant functions which is in contrast to the nature of FEs. Unfortunately, straightforward analytical evaluation of FE matrices becomes impossible when complexities such as curvature and continuous changes in material properties are introduced to problem assumptions. In chapter 5, a new universal matrix approach for evaluation of FE matrices is introduced. The approach is validated on a model Luneburg lens problem and shows perfect compatibility with the physics of the lens. There are a number of advantages for the approach among which here I suffice to mention the ease of universal coding and improved flexibility for multi-physical problems.